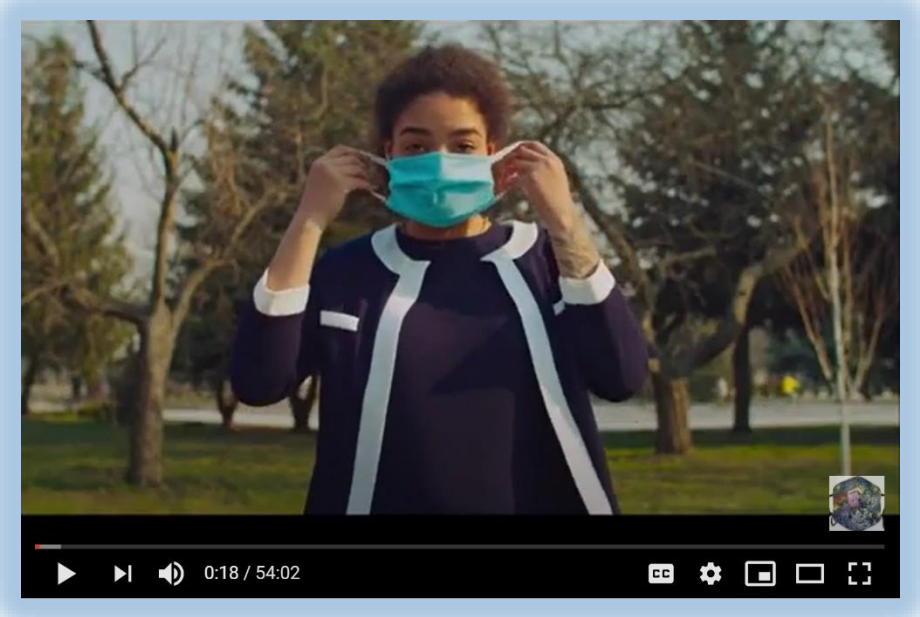
Want more info from The Science of Masks video?



[Attached are 50 PDFs of sources identified below.]

https://themodelhealthshow.com/maskfacts/ [Source]

https://www.fbcoverup.com/docs/library/2020-08-15-Mask-Facts-The-Science-and-History-of-Mask-in-Medicine-by-Shawn-Stevenson-The-Model-Health-Show-Aug-15-2020.mp4

Also on:

https://youtu.be/NxwRNmqo1mg

See the facts…

CLINICAL EVIDENCE & SUPPLEMENTARY DATA

Demonstrating ineffectiveness of masks in surgery

Disposable surgical face masks for preventing surgical wound infection in clean surgery.
Cochrane Systematic Review

READ THE ARTICLE

The evolution of the surgical mask: filtering efficiency versus effectiveness.
Infection Control & Hospital Epidemiology

READ THE ARTICLE

Surgical face masks in modern operating rooms—a costly and unnecessary ritual?
The Journal of Hospital Infection

READ THE ARTICLE

Wearing of Caps & Masks Not Necessary During Cardiac Catheterization.
Catheterization & Cardiovascular Diagnosis

READ THE ARTICLE

Masks-for-all for COVID-19 not based on sound data.
Center for Infectious Disease Research & Policy

READ THE ARTICLE

One Virus Particle is Enough to Cause Infectious Disease
Science Daily

READ THE ARTICLE

Measles Virus Can Live for up to Two Hours in an Airspace
CDC – US Gov

READ THE ARTICLE

Droplets and Aerosols in the Transmission of SARS-CoV-2.
The New England Journal of Medicine

READ THE ARTICLE

Closed environments facilitate secondary transmission of coronavirus

BMJ

READ THE ARTICLE

CLINICAL EVIDENCE & SUPPLEMENTARY DATA

Demonstrating ineffectiveness of masks in reducing the spread of infectious disease

A cluster randomised trial of cloth masks compared with medical masks in healthcare workers.

BMJ Open

READ THE ARTICLE

Unmasking the surgeons: the evidence base behind the use of facemasks in surgery.

Journal of the Royal Society of Medicine

READ THE ARTICLE

Surgical mask to prevent influenza transmission in households: a cluster randomized trial.

PLoS One

READ THE ARTICLE

Use of surgical face masks to reduce the incidence of the common cold.

American Journal of Infection Control

READ THE ARTICLE

Systematic review of the efficacy of face masks and respirators against coronaviruses and other respiratory transmissible viruses for the community, healthcare workers and sick patients.

The International Journal of Nursing Studies

READ THE ARTICLE

N95 masks v. SARS-CoV-2 particle size

American Chemical Society

READ THE ARTICLE

Surgical Mask vs N95 Respirator for Preventing Influenza Among Health Care Workers.
JAMA

READ THE ARTICLE

Effectiveness of N95 respirators versus surgical masks in protecting health care workers from acute respiratory infection
The Canadian Medical Association Journal

READ THE ARTICLE

The efficacy of medical masks and respirators against respiratory infection in healthcare workers.
Influenza & Other Respiratory Viruses

READ THE ARTICLE

Popular Mask Effectiveness Studies That Ignore Real World Conditions

(And Demonstrate Even More Mask Ineffectiveness Upon Further Investigation)

Frequently referenced hamster study with the dubious headline, "Surgical Masks Can Reduce Spread of Covid-19 virus by up to 75%". To replicate "real-life situations" the lead researcher placed surgical masks between the cages of hamsters in an isolated facility. If this study replicates real-life situations to the researcher, we should all be seriously concerned by the type of social life he has.

Hong Kong University

READ ARTICLE NOW

Researchers at Texas A&M declare that face masks prevented more than 66,000 infections in New York City in less than a month. The study makes hasty assumptions of universal consistency in mask quality, mask fit, duration of time worn, assumptions that the wearers aren't touching their face, that their aerosols and droplets aren't being deflected and sprayed all over their face, hair, clothes, and creating clouds of droplets from above, below, and through the sides of the mask. And it's also assuming that the masks are effective in the first place, which the vast majority of real-world clinical trials show that they're not. Lastly, their graph demonstrates a downward trend in infections prior to mandatory masking and doesn't differentiate the effects from social distancing, shelter-in-place, etc.

PNAS

READ ARTICLE NOW

Sexual contact carries some risk for exposure to infection with severe acute respiratory syndrome coronavirus 2 during the coronavirus disease 2019 pandemic.

Annals of Internal Medicine

READ ARTICLE NOW

Circulated articles state things like, "a cloth mask offers more protection than a surgical mask for people nearby." With data gathered from participants putting a mask on for a few moments and coughing 5 times to establish said effectiveness. Upon further review, the study itself noted, "both surgical and

cotton masks seem to be ineffective in preventing the dissemination of SARS–CoV-2 from the coughs of patients with COVID-19 to the environment." Even though they came to that conclusion with what can easily be considered an inadequate amount of data (they've since retracted), what's even more alarming is that the scientists found when test subjects coughed into the masks, even more virus particles ended up on the OUTSIDE of the mask than on the inside of the mask. It may come as a surprise, but that's simply how viruses can travel in the real world.

Annals of Internal Medicine

READ ARTICLE NOW

Yet again, another study positing the efficacy of masks by having participants breathe/cough into a collection apparatus. This time it was a G-II bioaerosol collecting device. To repeat, this on and off duration of mask use and direct collection of particles neglects how viruses travel in the real world and how the duration of mask usage makes them exceedingly less effective.

Nature Medicine

READ ARTICLE NOW

CLINICAL EVIDENCE & SUPPLEMENTARY DATA

Demonstrating the physiological and psychological damage caused by mask usage

The physiological impact of wearing an N95 mask during hemodialysis.

Journal of the Formosan Medical Association

READ THE ARTICLE

Respiratory consequences of N95-type Mask usage in pregnant healthcare workers.

Antimicrobial Resistance and Infection Control

READ THE ARTICLE

Carbon dioxide re-breathing with close fitting face respirator masks.

Anaesthesia

READ THE ARTICLE

Carbon dioxide rebreathing in respiratory protective devices.

Ergonomics

READ THE ARTICLE

Surgical mask induced deoxygenation.

Neurocirugia

READ THE ARTICLE

Protective Facemask Impact on Human Thermoregulation.

The Annals of Occupational Hygiene

READ THE ARTICLE

Amygdala Responsivity to High-Level Social Information from Unseen Faces.

The Journal of Neuroscience

READ THE ARTICLE

Rise in childhood mental health issues resulting from the social lockdown.

University of Oxford

READ THE ARTICLE

Mental health issues in children amidst COVID-19 pandemic.

The Canadian Medical Association Journal

READ THE ARTICLE

Effects of wearing N95 and surgical facemasks on heart rate, thermal stress and subjective sensations.

The International Archives of Occupational and Environmental Health

READ THE ARTICLE

Immune cells become overactive when oxygen levels are deranged.

University of Edinburgh

READ THE ARTICLE

Stress and the Human Immune System.

Journal of Evolutionary Biology – Correction: Psychological Bulletin

READ THE ARTICLE

The Developmental Origins of the Social Brain.

Frontiers in Psychology

READ THE ARTICLE

CLINICAL EVIDENCE & SUPPLEMENTARY DATA

Demonstrating health/immune system/covid-19 relationship and coronavirus susceptibility

Up to 650,000 people die of respiratory diseases linked to seasonal flu each year.
World Health Organization

READ THE ARTICLE

SARS-CoV-2 mutation
Cell

READ THE ARTICLE

Comorbidities the rule in New York's COVID-19 deaths.
NY State Dept. of Health

READ THE ARTICLE

Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area.
JAMA

READ THE ARTICLE

The State of US Health Burden of Diseases.
JAMA

READ THE ARTICLE

Spotlight on World Obesity Rates.
Central Intelligence Agency (CIA)

READ THE ARTICLE

FDA clears IND application for natural killer cell-based COVID-19 therapy.
Healio

READ THE ARTICLE

People with low NK cell counts had much higher rates of severe infections from COVID-19.

Shawn Stevenson. (Aug. 15, 2020). Mask Facts: "The Science & History of Mask in Medicine." The Model Health Show.

Medical Hypothesis

READ THE ARTICLE

Immune response to a 30-minute walk.

Medicine & Science in Sports & Exercise

READ THE ARTICLE

Sleep-deprivation reduces NK cell number and function.

Psychoneuroendocrinology

READ THE ARTICLE

The impact of stressful life events on natural killer cells.

Stress & Health

READ THE ARTICLE

Sleep deprivation effect on the immune system mirrors physical stress.

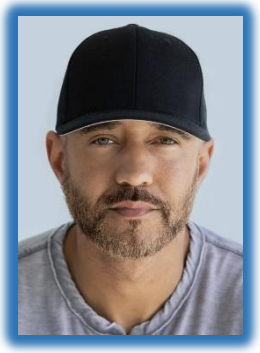
American Academy of Sleep Medicine

READ THE ARTICLE

The prevalence of stress in primary care.

JAMA Internal Medicine

READ THE ARTICLE



These resources are packed with real world, clinically proven data too are informed and empowered

About the Author & Producer: Shawn Stevenson

I Received A Wake-Up Call From Within

My health was failing, my relationships were failing, I was struggling in school, and I was broke and broken. But when things were darkest, I found a glimmer of light. While sitting on my bed one night, about to down my prescription and over-the-counter drugs to help me sleep through my pain, my grandmother came rushing into my mind. She had been calling and checking on me throughout this 2-year process, but I would always act like things were alright. Though, clearly, they were not alright. From there I began writing books and creating podcasts through The Model Health Show. Outside of my family, it's become the real love of my life to be of service this way. The show now reaches millions of listeners each year, and my passion continues to grow each day.

Cochrane Library

Cochrane Database of Systematic Reviews

Disposable surgical face masks for preventing surgical wound infection in clean surgery (Review)

Vincent M, Edwards P

Vincent M, Edwards P.
Disposable surgical face masks for preventing surgical wound infection in clean surgery.
Cochrane Database of Systematic Reviews 2016, Issue 4. Art. No.: CD002929.
DOI: 10.1002/14651858.CD002929.pub3.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Disposable surgical face masks for preventing surgical wound infection in clean surgery

Marina Vincent[1], Peggy Edwards[2]

[1]Department of Health Sciences, University of York, York, UK. [2]C/o Cochrane Wounds, School of Nursing, Midwifery and Social Work, University of Manchester, Manchester, UK

Contact address: Marina Vincent, Department of Health Sciences, University of York, Heslington, York, YO10 5DD, UK. marina.vincent7@gmail.com, marina.vincent7@gmail.com.

Editorial group: Cochrane Wounds Group.
Publication status and date: New search for studies and content updated (no change to conclusions), published in Issue 4, 2016.

Citation: Vincent M, Edwards P. Disposable surgical face masks for preventing surgical wound infection in clean surgery. *Cochrane Database of Systematic Reviews* 2016, Issue 4. Art. No.: CD002929. DOI: 10.1002/14651858.CD002929.pub3.

ABSTRACT

Background

Surgical face masks were originally developed to contain and filter droplets containing microorganisms expelled from the mouth and nasopharynx of healthcare workers during surgery, thereby providing protection for the patient. However, there are several ways in which surgical face masks could potentially contribute to contamination of the surgical wound, e.g. by incorrect wear or by leaking air from the side of the mask due to poor string tension.

Objectives

To determine whether the wearing of disposable surgical face masks by the surgical team during clean surgery reduces postoperative surgical wound infection.

Search methods

In December 2015, for this seventh update, we searched: The Cochrane Wounds Specialised Register; The Cochrane Central Register of Controlled Trials; Ovid MEDLINE; Ovid MEDLINE (In-Process & Other Non-Indexed Citations); Ovid EMBASE and EBSCO CINAHL. We also searched the bibliographies of all retrieved and relevant publications. There were no restrictions with respect to language, date of publication or study setting.

Selection criteria

Randomised controlled trials (RCTs) and quasi-randomised controlled trials comparing the use of disposable surgical masks with the use of no mask.

Data collection and analysis

Two review authors extracted data independently.

Main results

We included three trials, involving a total of 2106 participants. There was no statistically significant difference in infection rates between the masked and unmasked group in any of the trials. We identified no new trials for this latest update.

Authors' conclusions

From the limited results it is unclear whether the wearing of surgical face masks by members of the surgical team has any impact on surgical wound infection rates for patients undergoing clean surgery.

PLAIN LANGUAGE SUMMARY

Disposable surgical face masks for preventing surgical wound infection in clean surgery

Background

Surgeons and nurses performing clean surgery wear disposable face masks. The purpose of face masks is thought to be two-fold: to prevent the passage of germs from the surgeon's nose and mouth into the patient's wound and to protect the surgeon's face from sprays and splashes from the patient. Face masks are thought to make wound infections after surgery less likely. However, incorrectly worn masks may increase the likelihood of the wound getting contaminated with germs. We wanted to discover whether wearing a face mask during surgery makes infections of the wound more likely after the operation.

Review question

This review aimed to find out if wearing disposable face masks increases or decreases the number of cases of wound infection after clean surgery.

Study characteristics

We searched for all studies that had been done in the past relevant to this topic. Studies included in our analysis were those looking at the use of face masks in 'clean' surgery in adults and children. Clean surgery is when the operation does not go into organs that may contain bugs such as the lungs, gut, genitals and bladder. Infections of the wound are less likely to occur after 'clean' surgery, compared to 'unclean' surgery. We chose to look at this type of surgery because infections occurring after clean surgery would more likely be due to the use of the face mask, and not because of the nature of the operation. We also only looked at one particular type of study, the randomised controlled trial (RCT), where the people involved (participants) were randomly put into one of two groups: one group where the surgical team wore a face mask during the operation and one group where the surgical team did not wear a face mask. We compared the number of wound infection cases occurring after surgery between two groups.

Key results

Overall, we found very few studies and identified no new trials for this latest update. We analysed a total of 2106 participants from the three studies we found. All three studies showed that wearing a face mask during surgery neither increases nor decreases the number of wound infections occurring after surgery. We conclude that there is no clear evidence that wearing disposable face masks affects the likelihood of wound infections developing after surgery.

Quality of the evidence

The findings from this review cannot be generalised for several reasons: the studies included only looked at clean surgery, some of the studies did not specify what type of face mask was used and one of the studies did not involve many participants therefore making the findings less credible. The quality of the studies we found was low overall. The way in which participants were selected for the studies was not always completely random, which means the authors' judgements could have influenced the results. More research in this field is needed before making further conclusions about the use of face masks in surgery.

This plain language summary is up to date as of 22nd December 2015.

BACKGROUND

Description of the condition

Surgical face masks were originally developed to contain and filter droplets containing microorganisms expelled from the mouth and nasopharynx during surgery. They were introduced around a century ago as a method of protecting patients from the risk of surgical wound infections (Belkin 1997). The costs incurred when a patient contracts a surgical wound infection are considerable in financial as well as social terms. It has been estimated that each patient with a surgical wound infection requires an additional hospital stay of 6.5 days and that hospital costs are doubled (Plowman 2000). When extrapolated to all acute hospitals in England, it is estimated that the annual cost nationally is almost GBP 1 billion.

Description of the intervention

The primary purpose of a surgical mask is to provide protection for the patient from the surgical team. Masks have also been advocated as a barrier to protect the surgical team from the patient (Garner 1996; Weber 1993). This systematic review does not investigate the use of surgical masks for this purpose.

Surgical face masks are disposable and generally made up of three or four layers, often with two filters that prevent passage of material greater than 1 micron, therefore trapping bacteria of that size or larger. Face masks of this type are claimed to provide protection for a minimum of four hours (UHS 2000). Worn correctly, the mask should cover the nose with the metal band contouring the bridge of the nose. The mask should be drawn underneath the mouth and secured by tying the tapes firmly around the back of the head.

Although the surgical mask is designed to protect the patient, there are several ways in which it could actually contribute to the contamination of surgical wounds. Firstly, insufficient tension on the strings causes 'venting', or leakage of air from the side of the mask. The exhalation of moist air increases resistance, which is thought to exacerbate the problem of venting (Belkin 1996). Secondly, Belkin 1996 also cites 'wicking' as a method of conveying liquid via capillary action as possibly contributing to the passage of bacteria. Thirdly, a mask could cause contamination by 'wiggling'. This is a term used to describe friction of the mask against the face, which has been shown to cause the dispersal of skin scales from the face resulting in possible contamination of surgical wounds (Schweizer 1976). In addition, the mask may be worn incorrectly, for example, allowing exposure of the nose or mouth. Removal of the mask by grasping the filter section could result in contamination of the wearer's hands whereas disposal is recommended by handling the tapes only (Perry 1994).

How the intervention might work

These issues call into question the effectiveness of the design and highlight the incorrect use of surgical face masks. As with many interventions, surgical face masks were introduced without standard specifications or formal evaluation. Despite acknowledging the controversy surrounding the use of masks, they are currently recommended by numerous operating department organisations (AORN 1998; AfPP 2007).

There is evidence that face mask practice is inconsistent, possibly due to an inadequate rationale for their use. For example, the use of surgical face masks has been abandoned by some surgical teams (in part or whole) and during certain procedures. In choosing to not wear a mask, members of the surgical team could be leaving the patient vulnerable to the risk of wound infection via droplet contamination.

A clean surgical wound is classified as "an uninfected operative wound in which no inflammation is encountered and the respiratory, alimentary, genital or uninfected urinary tract is not entered" (Mangram 1999). Non-clean wounds may be classified as clean-contaminated, contaminated or dirty-infected, depending upon the area of the body operated upon and the level of infection and inflammation present. A surgical wound is less likely to become infected postoperatively if it is classified as clean, therefore any infection arising could be more reasonably attributed to other factors such as the use of a surgical face mask (Mangram 1999).

Diagnosis of a surgical wound infection is not without its challenges. For example, some patients such as the elderly and the immunocompromised do not always display the cardinal signs of infection. However, correct diagnosis of surgical wound infections is imperative to ensure accurate surveillance. A surgical wound infection is defined by purulent drainage and at least one of the following signs or symptoms: pain, localised swelling, redness or heat (Mangram 1999).

Why it is important to do this review

The above discussion indicates that the role of the surgical mask as an effective measure in preventing surgical wound infections is questionable and warrants a systematic review.

OBJECTIVES

To determine whether the wearing of disposable surgical face masks by the surgical team during clean surgery reduces postoperative surgical wound infection.

METHODS

Criteria for considering studies for this review

Types of studies

Randomised controlled trials (RCTs) and quasi-randomised controlled trials comparing the use, by members of the surgical team, of disposable surgical masks with the use of no mask.

Types of participants

Adults and children undergoing clean surgery.

Types of interventions

The specific comparison to be made is the wearing, by the surgical team (scrubbed and not scrubbed), of disposable surgical face masks compared with no masks. Due to the difference in specifications, we used the trial author's definition of disposable surgical mask.

Types of outcome measures

Primary outcomes

- The incidence of postoperative surgical wound infection (the definition of wound infection used by the trial authors is used throughout).

Secondary outcomes

- Costs.
- Length of hospital stay.
- Mortality rate.

Publication date, language and publication status did not influence eligibility decisions.

Search methods for identification of studies

Electronic searches

For this seventh update, we searched the following databases to identify reports of relevant clinical trials:

- The Cochrane Wounds Specialised Register (searched 22 December 2015);
- The Cochrane Central Register of Controlled Trials (CENTRAL) (The Cochrane Library 2015, Issue 11);
- Ovid MEDLINE (1946 to 22 December 2015);
- Ovid MEDLINE (In-Process & Other Non-Indexed Citations) (searched 22 December 2015);
- Ovid EMBASE (1974 to 22 December 2015);
- EBSCO CINAHL Plus (1937 to 23 December 2015).

The search strategies used for these databases can be found Appendix 1. We combined the Ovid MEDLINE search with the Cochrane Highly Sensitive Search Strategy for identifying randomised trials in MEDLINE: sensitivity- and precision-maximising version; Ovid format (Lefebvre 2011). We combined the EMBASE search with the Ovid EMBASE trial filter terms developed by the UK Cochrane Centre (Lefebvre 2011). We combined the CINAHL searches with the trial filter terms developed by the Scottish Intercollegiate Guidelines Network (SIGN 2015).There were no restrictions with respect to language, date of publication or study setting.

Searching other resources

We searched the bibliographies of all retrieved and relevant publications identified by these strategies for further studies.

Data collection and analysis

Selection of studies

Two review authors independently assessed titles and abstracts of references identified by the search strategy according to the selection criteria. We obtained copies of those articles and studies that appeared to satisfy these criteria in full. When it was unclear from the title or abstract if the paper fulfilled the criteria, or when there was disparity between the review authors, we obtained a full-text copy. The two review authors jointly decided whether the study met the inclusion criteria. For this update, one review author assessed titles and abstracts of references identified by the search strategy. Again, when it was unclear from the title or abstract if the study fulfilled the criteria, the full-text was obtained and reviewed by one review author, all decisions were discussed with a member of the editorial team of Cochrane Wounds.

Data extraction and management

We used a piloted data extraction sheet to extract and summarise details of the studies. When data were missing from the study, we attempted to contact the trial authors to obtain missing information. Data extraction was undertaken independently by the two review authors and compared. We excluded studies if they were not randomised or quasi-randomised trials of disposable surgical face masks. Excluded studies are listed in the Characteristics of excluded studies table with reasons for their exclusion.

We extracted the following data from each study.

- Trial setting.
- Number of air filtration changes in the surgical field per hour.
- Filtering capacity/specification of masks.
- Types of surgery.
- Number of wound infections.
- Definition of wound infection.
- Depth of wound infection.
- Documentation of co-interventions.
- Use of prophylactic antibiotics.
- Use of antiseptic irrigation.
- Identified bacteria associated with staff and patients.
- Measurement of compliance in the wearing of surgical face masks (i.e. mask covered nose and mouth, presence of wicking and venting).
- The size of the surgical team.

Assessment of risk of bias in included studies

Two review authors independently assessed each included study using the Cochrane tool for assessing risk of bias (Higgins 2011). This tool addresses six specific domains, namely sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting and other issues (e.g. extreme baseline imbalance) (see Appendix 2 for details of the criteria on which each judgement was based). We assessed the studies to detect potential sources of bias in the study design. We extracted data regarding the following aspects of risk of bias.

- Method of randomisation: how the randomisation schedule was generated, the method of randomisation, e.g. envelopes, computer etc.
- Allocation concealment.
- Blinding of patients (recipients).
- Blinding of outcome assessors to wearing of masks.
- Extent of loss to follow-up and use of intention-to-treat analysis.
- Source of funding.
- Early stopping.
- Baseline comparability of treatment and control groups.

Data synthesis

We entered data into the Cochrane Review Manager (RevMan) software (RevMan 2014). Results are presented with 95% confidence intervals (CI). Methods of synthesising studies were dependent upon the quality, design and heterogeneity of the studies identified. We reported estimates for dichotomous outcomes as odds ratio (OR) as the event rate was less than 30% (Altman 1991). Where synthesis was inappropriate, we undertook a narrative overview.

RESULTS

Description of studies

Results of the search

The initial search, for the original review, yielded 250 citations; we examined the abstracts of these papers to assess potential relevance. We subsequently retrieved 97 papers for fuller examination. Of these, 84 were clearly not relevant to the review and 13 appeared potentially relevant. We subsequently excluded 11 from the review due to study design, or ineligible outcome measures (e.g. bacterial load). We included two studies. We identified no unpublished studies that met the criteria for inclusion. There was no response to requests for further information from the authors of two included studies (Chamberlain 1984; Tunevall 1991). No studies were published in duplicate. During subsequent updates of the review, we identified five further studies; four did not meet the inclusion criteria after assessment (Alwitry 2002; McGovern 2013; Salassa 2014; Sjol 2002), and one met the criteria for inclusion and we added it to the review (Webster 2010). We identified no new trials for this latest update.

This review took at face value any description in the original studies of the type and cleanliness category of surgery performed. In one study, we contacted the author who provided data for clean surgery only (Webster 2010). As a result, we included studies performed in the operating department and excluded other areas such as the laboratory, maternity ward and accident and emergency.

Included studies

See the Characteristics of included studies table.

Type of surgery

Tunevall 1991 included all types of surgery: clean, clean-contaminated and contaminated. Chamberlain 1984 involved gynaecological operation lists carried out by masked and unmasked staff. Webster 2010 randomised non-scrubbed staff per list into masked and unmasked groups. Surgery included obstetrics, gynaecology, general, orthopaedics, breast and urological. We only extracted data relating to clean surgery from all three studies.

Type of mask

Only one study specified the types of face mask used (Tunevall 1991), which were Comfort Clinimask (Molnycke), Surgine II antifog mask (Surgikos) and Aseptex (3M). In one study the type of mask was not mentioned (Chamberlain 1984), and in the other study standard masks were used (Webster 2010).

Number of patients

A power calculation informed Tunevall 1991 that their study would have to include over 3000 patients to demonstrate a decrease of 30% in the wound infection rate. It is unclear whether the power calculation took account of the clustered nature of the data. Although the Tunevall's study involved a total of 3088 patients, only 1429 patients undergoing clean surgery met the criteria for this review. In the study by Chamberlain 1984 only 41 patients were recruited because the study was discontinued. Out of this number, only 24 cases were clean surgery. With such a small number of female patients in this study, it is unlikely that they were representative of the population. Webster 2010 calculated that a sample size of at least 450 in each arm of the study would be needed to detect a 40% difference in surgical site infection rate between the two groups. Although 827 enrolled on the study, only 653 patients undergoing clean surgery met the criteria for this review (communication with trial author).

Outcome measures

The outcome measure used in Tunevall 1991 was wound infection defined as pus visible to the naked eye, or cellulitis without pus, both requiring debridement or percutaneous drainage and/or antibiotic therapy. With this study, follow-up was until after discharge but it was not explicit how these patients were followed up once discharged. Chamberlain 1984 did not define wound infection, but two out of the three wound infections reported were noted as serious enough to warrant antibiotics, the other infection being identified by a high vaginal swab. All patients in this study were examined daily until discharge. Webster 2010 used the National Nosocomial Infection Surveillance system, which categorises surgical site infections as superficial incisional, deep incisional and organ space. Follow-up was up to six weeks with the mean being 33.4 days for both groups.

None of the studies took any steps to measure compliance in relation to the correct wearing of surgical face masks, or recorded any events such as venting, wicking or wiggling. No study considered the other secondary outcome measures listed in this review.

Consent

One study author specified that consent was obtained from the staff involved in the study (Webster 2010). Tunevall 1991 stated that consent was obtained from patients, but Chamberlain 1984 and Webster 2010 did not specify that consent from patients had been obtained.

Excluded studies

We added a total of 15 studies to the Characteristics of excluded studies table. In summary, we excluded six studies because the focus of the study was not on assessing the rate of surgical site infection (Alwitry 2002; Ha'eri 1980; McGovern 2013; Norman 1995; Ritter 1975; Tunevall 1991). We excluded two studies because variables in addition to the rate of surgical site infection and the use of face masks were investigated (Berger 1993; Ruthman 1984). We excluded three studies because they did not involve any surgery and, rather, were simulation-based (Hubble 1996; McLure 1998; Mitchell 1991). Two studies were not RCTs or quasi-RCTs (Salassa 2014; Sjol 2002), one study assessed surgical site infection through the means of a patient questionnaire (Moore 2001), and one study did not state how many clean operations were included in their study (Orr 1981).

Risk of bias in included studies

See Figure 1 for the graph showing the review author's judgements about each 'Risk of bias' item presented as percentages across all included studies. See also Figure 2 for the summary showing the review author's judgements about each 'Risk of bias' item

Figure 1. Methodological quality graph: review authors' judgements about each methodological quality item presented as percentages across all included studies.

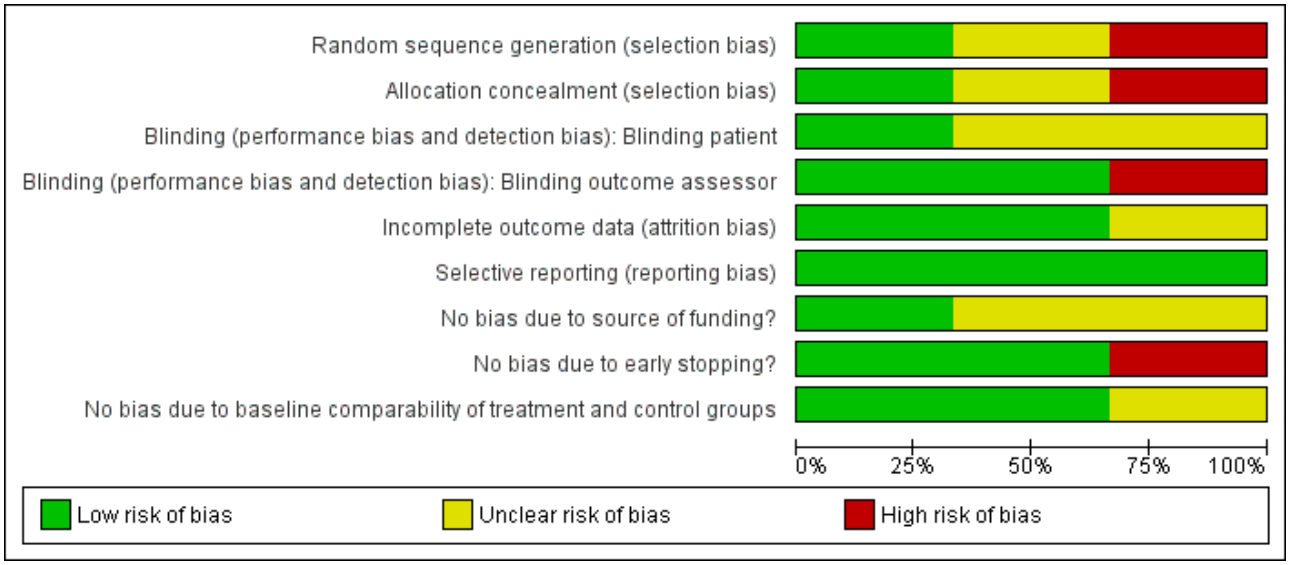
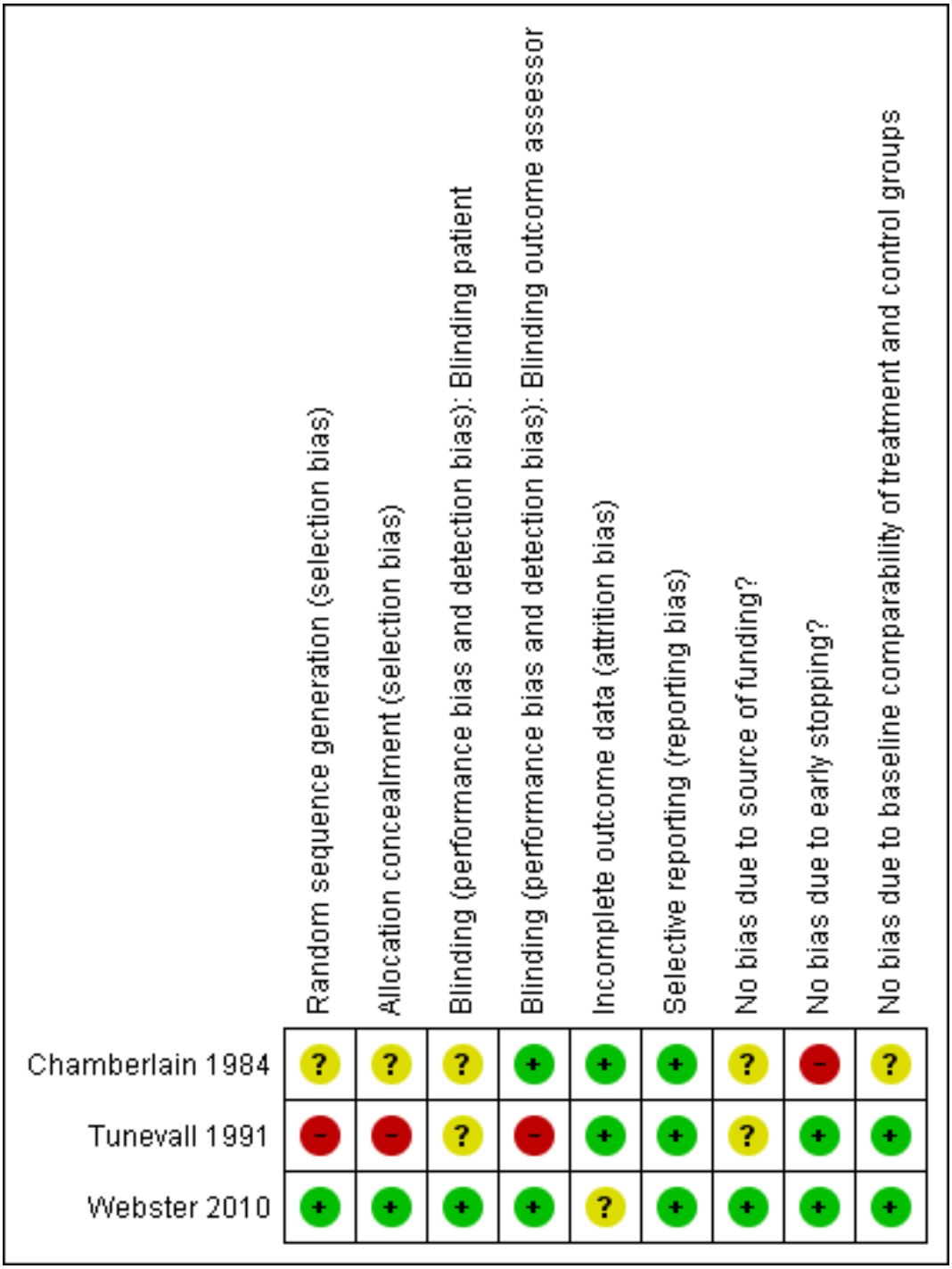


Figure 2. Methodological quality summary: review authors' judgements about each methodological quality item for each included study.



Allocation

Neither Chamberlain 1984 nor Tunevall 1991 used true randomisation with allocation concealment. Tunevall 1991 set up a random list for one year at a time denoting weeks as masked or unmasked but did not describe the method by which weeks were randomised to be masked/unmasked. A week, rather than an operating list or single operation, was the unit of allocation chosen for a period of one year, to ensure a similar number of major and minor cases (most major cases were performed at the beginning of the week). The randomisation list was inversed for the second and part of the third year due to anticipated seasonal differences. Allocation was not concealed as members of the theatre team were able to calculate whether any week was likely to be masked or unmasked. It is not clear whether the members of the admitting personnel had access to the randomisation list.

Chamberlain 1984 stated that patients on the operating lists of one surgical team were randomly allocated to a masked or unmasked group over two months. Later he indicated that masked and unmasked staff carried out the gynaecological operation lists alternately. The time between allocation of each list as masked or unmasked and the start of the list is not stated, making the extent of allocation concealment unclear.

Webster 2010 randomised participants per operating list. Allocation was concealed as randomisation occurred immediately before the start of the operating list via a phone call to a person blinded to the type of list.

In all studies the surgical team was the unit of randomisation and the patient was the unit of assessment, thus creating a unit of analysis error. There is no information in any study as to how patients were allocated to particular operating lists and so selection bias cannot be excluded.

Blinding

It was impossible to blind the care providers of the trials to wearing or omitting a surgical face mask. The blinding of patients was described by Webster 2010 but not by either Chamberlain 1984 or Tunevall 1991. No study distinguished between the use of local anaesthetic and general anaesthetic. Blinding of outcome assessors was achieved for Chamberlain 1984, where members of laboratory staff were unaware of the group allocation of the specimens obtained. Outcome assessors were also blinded in Webster 2010, where details of surgical site infections were obtained via routine surveillance or staff blinded to the intervention. In Tunevall 1991, specific notification of the trial was given with each wound swab submitted for culture, allowing the potential for detection bias.

Two studies included all members of the surgical team and neither of those studies examined whether particular members of the team were more or less likely to cause a surgical wound infection (Chamberlain 1984; Tunevall 1991). One study included only non-scrubbed staff (Webster 2010).

Incomplete outcome data

Chamberlain 1984 and Tunevall 1991 did not undertake an intention-to-treat analysis. Webster 2010 performed an intention-to-treat analysis. Chamberlain 1984 was discontinued after seven weeks after a third case of postoperative infection in the unmasked group was diagnosed. However the trial authors acknowledged that, although two of three wounds grew *Staphylococcus aureus*, in neither case was it a strain that corresponded to those isolated from the staff. No drop-outs were reported in Tunevall 1991. Webster 2010 reported seven drop-outs for clean surgery.

Other potential sources of bias

Source of funding

Two studies did not state a source of funding (Chamberlain 1984; Tunevall 1991), and one study declared a grant from Queensland Health Nursing Research (Webster 2010).

Early stopping of trial

Chamberlain 1984 was discontinued after seven weeks after a third case of postoperative infection in the unmasked group was diagnosed; this may well have been a chance difference, so potentially biasing the results in favour of masking.

Baseline imbalance

A description of the baseline characteristics of the patients is important to decide whether the results are generalisable and to compare characteristics of the two groups to ensure that the randomisation was successful. Tunevall 1991 confirmed baseline comparability for age and types of surgery. All patients in Chamberlain 1984 were female undergoing gynaecological surgery; no baseline comparability was reported. Groups were similar at baseline in Webster 2010 in terms of surgery, wound and American Society of Anaesthesiologists (ASA) classification as well as age, gender, preoperative hospitalisation, weight and prophylactic antibiotics.

Effects of interventions

The included studies compared the use of disposable surgical face masks with using no surgical face masks. A total of 2106 patients, undergoing clean surgery, were included in this review. We assessed clinical and methodological homogeneity. The observed clinical heterogeneity between the trials was reflected in parameters such as study population, time lapse between the first and latest study influencing technique and equipment, diagnosis and length of follow-up. Potential sources of clinical heterogeneity could be attributed to type of disposable surgical face mask, restricting non-scrubbed staff to the intervention group, operating theatre design (e.g. air flow rates) and country of study. Given this clinical heterogeneity, it was inappropriate to pool any of the studies.

Primary outcome: incidence of postoperative surgical wound infection

There were 2106 participants in three trials. Tunevall 1991 reported 13/706 (1.8%) postoperative wound infections in the masked group and 10/723 (1.4%) in the non-masked group (no statistically significant difference: odds ratio (OR) 1.34, 95% confidence interval (CI) 0.58 to 3.07). Chamberlain 1984 reported no postoperative wound infections in the masked group and 3/10 (30%) in the non-masked group (no statistically significant difference: OR 0.07, 95% CI 0.00 to 1.63). Webster 2010 reported 33/313 (10.5%) in the masked group and 31/340 (9.1%) in the non-masked group (no statistically significant difference: OR 1.17, 95% CI 0.70 to 1.97) (Analysis 1.1).

Secondary outcomes

None of the studies considered the secondary outcome measures specified in the review, i.e. costs, length of hospital stay and mortality rate.

DISCUSSION

Given the widespread use of surgical face masks, research into this topic remains surprisingly neglected. It was disappointing that only two trials met the inclusion criteria for the original review and these were undertaken prior to 1991. The inclusion of a more recent trial has helped to address the lack of evidence (Webster 2010).

Much of current national and international policy is based upon equivocal evidence from laboratory studies of the filtration efficiency of surgical face masks and of potential contamination of

the surgical field using settle plates. Such indirect evidence is of questionable clinical relevance.

Potential biases in the primary studies and the limitations they place on inferences

The strength of the evidence provided by the three studies that met the inclusion criteria for this review was weak. Two studies were quasi-randomised with unclear allocation concealment.

Methodologically, the results of Chamberlain 1984 and Tunevall 1991 may have been biased in several ways. Chamberlain 1984 did not specify the criteria used to detect the presence of a wound infection. Mangram 1999 reports that failure to use objective criteria to define surgical site infection has been shown to substantially affect reported surgical site infection rates. Chamberlain 1984 was limited by the discontinuation of the trial after seven weeks as result of several infections, thus creating a potential bias in the findings towards the use of surgical face masks.

Follow-up in Chamberlain 1984 continued until after discharge and up to discharge in Tunevall 1991. However the actual duration of follow-up could have varied considerably depending upon the type of surgery performed, with the potential for underestimating the number of surgical wound infections. Follow-up in Webster 2010 was more in keeping with international guidance of 30 days, but in some cases was less. It is likely that the inadequate allocation concealment and lack of blinding in the Chamberlain 1984 and Tunevall 1991 studies could have resulted in under or over-estimation of the effects of wearing a surgical face mask.

We were surprised at the small number of published studies. This could be due to a reluctance on the part of researchers to submit an equivocal trial for publication, and in turn for it to be accepted for publication. However, publication bias could not be tested by a funnel plot due to the small number of included studies.

Potential biases in the review and the limitations it places on inferences

We relied on the goodwill of experts in the field to provide information on completed or ongoing, published or unpublished studies. When critically appraising the validity of the studies we had to rely on adequate reporting of the trials. When there is minimal information in the trial report one cannot automatically assume that rigorous methods have not been followed. We attempted to obtain additional clarifying data from the investigators of two studies, however no responses were received. Webster 2010 provided data on patients undergoing clean surgery.

The examination of the effectiveness of disposable surgical face masks must be seen in the context of the number of variables associated with wound infections. It is difficult to interpret from small studies, such as Chamberlain 1984, whether the wearing of surgical face masks has an impact on rates of surgical wound infections in patients undergoing clean surgery.

Applicability of results

The results extracted for this review were limited to clean surgery and therefore cannot be extrapolated to other categories of surgery. The contribution that disposable surgical face masks make towards preventing infection is likely to be less consequential in contaminated wounds than in clean surgery.

The types of disposable surgical face mask used in the study were specified by Tunevall 1991 but not by Chamberlain 1984 or Webster 2010. It is possible that the specific mask composition changed in the years spanning the studies and this has the potential to influence results.

Although the review did not exclude trials involving the implantation of prostheses, we found no trials of this nature therefore limiting application of the review's results to this type of surgery. One study, Webster 2010, differentiated between scrubbed and non-scrubbed members of the team but, because only non-scrubbed staff were randomised into the study, it was not possible to discriminate between the contribution of the scrubbed and non-scrubbed members of the surgical team to any resulting surgical wound infection. It could be argued that non-scrubbed members of the team are less likely to be in a position to contaminate the surgical site.

All included studies were based in the operating department and so application of the results to other invasive procedures in other clinical areas is limited.

We examined the potential for surgical face masks to benefit the patient by reducing surgical wound infections or to harm the patient by increasing surgical wound infections in this review. We did not undertake analysis of the potential to harm or benefit the surgical team by way of protection. Although Chamberlain 1984 favoured the use of surgical face masks, the trial was relatively small and was discontinued due to the identification of wound infections in three out of the five major clean cases performed. This may have been a chance finding and thus these results are potentially biased in favour of wearing masks. Tunevall 1991 and Webster 2010 were larger trials, more rigorously designed and did not detect differences in the infection rate.

Both national and international guidelines acknowledge the controversy surrounding the use of disposable surgical face masks and yet continue to recommend their use. We found no other reviews in this area and the limited number of trials in this review make it unsafe to draw definitive conclusions about the effect of surgical face masks on reducing surgical wound infection in clean surgery.

AUTHORS' CONCLUSIONS

Implications for practice

From the limited results, it is unclear whether the wearing of surgical face masks by the surgical team either increases or reduces the risk of surgical site infection in patients undergoing clean surgery.

Implications for research

Important messages for future research:

1. The CONSORT statement should be used as a guideline for reporting of future trials (Schulz 2010).
2. Trials should be large enough to detect clinically important differences in infection rates.
3. Trials must discriminate between scrubbed and non-scrubbed personnel.

4. Trials must include clear definitions of surgery, surgical face masks and surgical wound infection.
5. Randomisation should be 'per operating list' (cluster randomisation) rather than 'per case' to avoid potential contamination of the surgical environment. To guard against selection bias, the randomisation allocation should be unpredictable, concealed and take place immediately prior to the commencement of the operating list.
6. Follow-up should be appropriate to the surgery performed. This may extend to the involvement of primary care.
7. Outcome assessors should be blinded to allocation.
8. Analysis should be by intention-to-treat of all patients following randomisation.
9. Economic evaluations should be incorporated into future trials.

Areas for further investigation include:

- disposable surgical face mask compared with wearing no mask;
- disposable surgical face mask compared with other mechanisms for protecting both patients and staff, such as visors/helmets.

ACKNOWLEDGEMENTS

Thanks to Alun Davies, Lorraine Foster, Michael Walker, Anne Humphreys and Pearl Shahib who refereed the protocol for relevance, readability and rigour. Thanks also to Vicky Arrowsmith, Sue Clark, Alun Davies, Michael Walker and Pearl Shahib who refereed the review and Jenny Bellorini who copy edited the recent updated review. We would like to thank Nicky Cullum and Andrea Nelson for methodological advice, Sally Bell-Syer for her support and the Cochrane Wounds Group for their help and assistance.

Allyson Lipp was the originator of this review and was responsible for the development of the protocol, the review and all updates to the present time. She has now retired and has stepped away from her author role. We would like to acknowledge her substantial involvement in this systematic review.

REFERENCES

References to studies included in this review

Chamberlain 1984 *{published data only}*

Chamberlain G, Houang E. Trial of the use of masks in the gynaecological operating theatre. *Annals of the Royal College of Surgeons of England* 1984;**66**(6):432-3.

Tunevall 1991 *{published data only}*

Tunevall T. Postoperative wound infections and surgical face masks: a controlled study. *World Journal of Surgery* 1991;**15**(3):383-8.

Webster 2010 *{published and unpublished data}*

Webster J, Croger S, Lister C, Doidge M, Terry M, Jones I. Use of face masks by non-scrubbed operating room staff: a randomized controlled trial. *ANZ Journal of Surgery* 2010;**80**(3):169-73.

References to studies excluded from this review

Alwitry 2002 *{published data only}*

Alwitry A, Jackson E, Chen H, Holden R. The use of surgical facemasks during cataract surgery: is it necessary?. *British Journal of Ophthalmology* 2002;**86**(9):975-7.

Berger 1993 *{published data only}*

Berger S, Kramer M, Nagar H, Finkelstein A, Frimmerman A, Miller H. Effect of surgical mask position on bacterial contamination of the operative field. *Journal of Hospital Infection* 1993;**23**:51-4.

Ha'eri 1980 *{published data only}*

Ha'eri G, Wiley A. The efficacy of standard surgical face masks: an investigation using 'tracer particles'. *Clinical Orthopaedics and Related Research* 1980;**148**:160-2.

Hubble 1996 *{published data only}*

Hubble M, Weale A, Perez J, Bowker K, MacGowen A, Bannister G. Clothing in laminar-flow operating theatres. *Journal of Hospital Infection* 1996;**32**(1):1-7.

McGovern 2013 *{published data only}*

McGovern PD, Albrecht M, Khan SK, Muller SD, Reed MR. The influence of surgical hoods and togas on airborne particle concentration at the surgical site: an experimental study. *Journal of Orthopaedic Science* 2013;**18**:1027-30.

McLure 1998 *{published data only}*

McLure H, Talboys C, Yentis S, Azadian B. Surgical face masks and downward dispersal of bacteria. *Anaesthesia* 1998;**53**(7):624-6.

Mitchell 1991 *{published data only}*

Mitchell N, Hunt S. Surgical face masks in modern operating rooms - a costly and unnecessary ritual?. *Journal of Hospital Infection* 1991;**18**:239-42.

Moore 2001 *{unpublished data only}*

Moore A, Klimach OE. Face masks or visors?. Unpublished study.

Norman 1995 *{published data only}*

Norman A. A comparison of face masks and visors for the scrub team. *British Journal of Theatre Nursing* 1995;**5**(2):10-3.

Orr 1981 *{published data only}*

Orr N. Is a mask necessary in the operating theatre?. *Annals of the Royal College of Surgeons of England* 1981;**63**:390-2.

Ritter 1975 *{published data only}*

Ritter M, Eitzen H, French M, Hart J. The operating room environment as affected by people and the surgical face mask. *Clinical Orthopaedics and Related Research* 1975;**111**:147-50.

Ruthman 1984 *{published data only}*

Ruthman J, Hendricksen D, Miller R, Quigg D. Effect of cap and mask on infection rates. *Illinois Medical Journal* 1984;**165**(6):397-9.

Salassa 2014 *{published data only}*

Salassa TE, Swiontkowski MF. Surgical attire and the operating room: role in infection prevention. *Journal of Bone and Joint Surgery* 2014;**96**:1485-92.

Sjol 2002 *{published data only}*

Sjol A, Kelbaek H. Is use of surgical caps and masks obsolete during percutaneous heart catheterization?. *Ugeskrift for Laeger* 2002;**164**(12):1673-5.

Tunevall 1992 *{published data only}*

Tunevall T, Jorbeck H. Influence of wearing masks on the density of airborne bacteria in the vicinity of the surgical wound. *European Journal of Surgery* 1992;**158**:263-6.

Additional references

AfPP 2007

Pirie S, Beesley J, Vernon T. Standards and Recommendations for Safe Perioperative Practice. England: Association for Perioperative Practitioners, 2007.

Altman 1991

Altman D. Practical Statistics for Medical Research. 1st Edition. London: Chapman Hall, 1991.

AORN 1998

American Operating Room Nurses. Standards, Recommendations Practice and Guidelines. Colorado: American Operating Room Nurses, 1998.

Belkin 1996

Belkin N. A century after their introduction, are surgical masks necessary?. *American Operating Room Nurses Journal* 1996;**64**(4):602-7.

Belkin 1997

Belkin N. The evolution of the surgical mask: filtering efficiency versus effectiveness. *Infection Control and Hospital Epidemiology* 1997;**18**(1):49-57.

Garner 1996

Garner JS. Guideline for isolation precautions in hospital. *Infection Control and Hospital Epidemiology* 1996;**17**(1):53-80.

Higgins 2011

Higgins JPT, Altman DG, on behalf of the Cochrane Statistical Methods Group and the Cochrane Bias Methods Group (Editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Lefebvre 2011

Lefebvre C, Manheimer E, Glanville J, on behalf of the Cochrane Information Retrieval Methods Group. Chapter 6: Searching for studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Mangram 1999

Mangram A, Horan T, Pearson M, Silver L, Jarvis W. Guideline for prevention of surgical site infection. *Infection Control and Hospital Epidemiology* 1999;**20**(4):247-78.

Perry 1994

Perry AG, Potter PA. Clinical Nursing Skills and Techniques. St Louis: Mosby, 1994.

Plowman 2000

Plowman R, Graves N, Griffin M. The Socio-economic Burden of Hospital Acquired Infection. London: Public Health Service Laboratory, 2000.

RevMan 2014 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Schulz 2010

Schulz K, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;**340**:c332.

Schweizer 1976

Schweizer RT. Mask wiggling as a potential cause of wound contamination. *Lancet* 1976;**2**:1129-30.

SIGN 2015

Scottish Intercollegiate Guidelines Network (SIGN). Search filters. http://www.sign.ac.uk/methodology/filters.html#random (accessed 06 April 2016).

UHS 2000

Universal Hospital Supplies. Personal correspondence 2000.

Weber 1993

Weber A, Willeke K, Marchioni R, Myojo T, McKay R, Donnelly J, et al. Aerosol penetration and leakage characteristics of masks used in the health care industry. *American Journal of Infection Control* 1993;**21**(4):167-73.

References to other published versions of this review

Lipp 2014

Lipp A, Edwards P. Disposable surgical facemasks for preventing surgical wound infection in clean surgery. *Cochrane Database of Systematic Reviews* 2014, Issue 2. [DOI: 10.1002/14651858.CD002929.pub2]

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Chamberlain 1984

Methods	Quasi-randomised controlled trial
Participants	41 female patients undergoing surgery; 24 clean and 17 non-clean Inclusion criteria: gynaecology Exclusion criteria: none stated Baseline comparability; none reported
Interventions	Group 1. Mask (n = 14) Group 2. No mask (n = 10)
Outcomes	Wound infection defined as serious enough to warrant antibiotics in 2 of the cases and via a high vaginal swab in the third case. Follow-up until discharge only.
Notes	Study discontinued due to 3 surgical wound infections in the unmasked group, although not proven as causal. Data extracted for clean surgery only. Unit of analysis error present.

Chamberlain 1984 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomly allocated per list, but method unclear
Allocation concealment (selection bias)	Unclear risk	Time between allocation of masked and unmasked list and the list start was unclear
Blinding (performance bias and detection bias) Blinding patient	Unclear risk	Not described
Blinding (performance bias and detection bias) Blinding outcome assessor	Low risk	Quote: "The laboratory work was carried out by a member of staff who was not aware of the group allocation of the specimens obtained." Comment: blinding of outcomes assessors reduces risk of performance and detection bias
Incomplete outcome data (attrition bias) All outcomes	Low risk	Intention-to-treat analysis not stated. No drop-outs reported.
Selective reporting (reporting bias)	Low risk	Prespecified outcomes reported on, but trial protocol not accessed
No bias due to source of funding?	Unclear risk	No funding sources stated
No bias due to early stopping?	High risk	The study was discontinued after the third case of postoperative infection in the unmasked group. The study authors state that the bacterial strain of the infections did not correspond to those isolated from the staff.
No bias due to baseline comparability of treatment and control groups	Unclear risk	Baseline comparability not stated. All participants were female undergoing gynaecological surgery.

Tunevall 1991

Methods	Quasi-randomised controlled trial
Participants	3088 patients undergoing general, vascular, breast, acute and elective surgery. Clean surgery was performed on 1429. Non-clean surgery was performed on 1659. Trial setting: operating department. Inclusion criteria: operation through intact skin and primary closure. Exclusion criteria: patients not informed or consent not given; outpatients; orthopaedics; urology; anal surgery; insertion of synthetic grafts; or haematologic disease. Baseline comparability: similar for age, acute and cold surgery.
Interventions	Group 1. Mask (n = 706) Group 2. No mask (n = 723)
Outcomes	Wound infection defined as visible pus and/or cellulitis without pus requiring debridement, drainage and/or antibiotics.

Tunevall 1991 *(Continued)*

	Duration of follow-up not stated but until after discharge from the ward.
Notes	Data extracted for clean surgery only. Patients had 2 to 3 body washes pre-operatively with 4% chlorhexidine prior to elective surgery. In most acute cases, at least one body wash was given. Unit of analysis error present.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quote: "A random list was set up for 1 year, denoting weeks as 'masked' or 'unmasked'. To avoid seasonal differences between the groups the list was inversed for the second and for the third part of the year." Comment: this makes selection at high risk of bias
Allocation concealment (selection bias)	High risk	Inadequate as investigators enrolling participants could possibly foresee allocation and thus introduce selection bias
Blinding (performance bias and detection bias) Blinding patient	Unclear risk	Not described
Blinding (performance bias and detection bias) Blinding outcome assessor	High risk	Notification of the trial was issued with each wound swab
Incomplete outcome data (attrition bias) All outcomes	Low risk	Not analysed on an intention-to-treat basis. No drop-outs reported.
Selective reporting (reporting bias)	Low risk	Prespecified outcomes reported on, but trial protocol not accessed
No bias due to source of funding?	Unclear risk	No funding sources stated
No bias due to early stopping?	Low risk	The trial was based on a power calculation and was not stopped early
No bias due to baseline comparability of treatment and control groups	Low risk	Baseline comparability stated for age and type of surgery

Webster 2010

Methods	Randomised controlled trial
Participants	811 patients undergoing gynaecological, obstetric, general (open), general (laparoscopic), urology and breast surgery. Clean surgery was performed on 660 patients and non-clean on 151 patients. Inclusion criteria: none stated Exclusion criteria: surgery where a mask was specifically required, e.g. air borne infection Participants were similar at baseline for age, gender, weight, prophylactic antibiotics and ASA classification

Webster 2010 *(Continued)*

Interventions	Group 1. Mask (n = 313) Group 2. No mask (n = 340)
Outcomes	Wound infection defined by criteria used by National Nosocomial Infection Surveillance System: superficial incisional, deep incisional and organ space Group 1. Mean follow-up 33.4 days (SD 22.1) Group 2. Mean follow-up 33.4 days (SD 22.8)
Notes	Missing data for 7 clean cases. Unit of analysis error present. Quote: "Only non-scrubbed staff, including anaesthetists, were asked to comply with the random assignment." Comment: scrubbed staff were not included in the trial

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Operating lists were randomised into two arms, mask group and no mask group using a computer-generated randomisation schedule." Comment: This precaution reduces the risk of selection bias.
Allocation concealment (selection bias)	Low risk	Quote: " Allocation occurred immediately before the commencement of the session, following a phone call to a person who was unaware of the type of list in each theatre". Comment: this precaution reduces the risk of selection bias
Blinding (performance bias and detection bias) Blinding patient	Low risk	Patients were unaware of treatment allocation
Blinding (performance bias and detection bias) Blinding outcome assessor	Low risk	Quote: "Details about any post operative wound infection was obtained by routine surveillance methods, that is by the medical officer, ward staff or infection control nurse who were blinded to the treatment protocol."
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Clean data not analysed on an intention-to-treat basis; 7 drop-outs reported
Selective reporting (reporting bias)	Low risk	Prespecified outcomes reported on, but trial protocol not accessed
No bias due to source of funding?	Low risk	Quote: "JW received grant support through two Queensland Health Nursing Research Grants." Comment: this grant is unlikely to have biased the results of the trial
No bias due to early stopping?	Low risk	The trial was based on a power calculation and was not stopped early
No bias due to baseline comparability of treatment and control groups	Low risk	Groups were comparable for baseline characteristics of type of surgery, wound and ASA classification as well as age, gender, preoperative hospitalisation, weight and prophylactic antibiotics

ASA: American Society of Anaesthesiologists
SD: standard deviation

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Alwitry 2002	The measurement of bacterial load was used rather than infection rates
Berger 1993	The study was concerned with both contamination and wound infection. It was poorly designed as all procedures had varying mask positions at different times of the procedure. It was impossible to distinguish from the results the masked and unmasked periods. Settle plates were used to measure contamination and no infections were recorded. This study was discontinued after recruitment of 30 patients due to the unacceptable level of contamination of the settle plates.
Ha'eri 1980	This study was primarily concerned with surgical site contamination by human albumen microspheres and not surgical wound infection
Hubble 1996	Excluded as it was a theatre-based simulation that did not involve any surgery. Contamination was measured using settle plates at various distances from the participant. This study included hats as well as masks in traditional and laminar flow theatres.
McGovern 2013	The effect of different surgical gowns on counts of airborne particles was investigated in this study, with the primary outcome being mean particle count (not rate of postoperative surgical wound infection).
McLure 1998	A laboratory simulation involving the analysis of bacterial colonies on agar plates. No surgery was involved.
Mitchell 1991	An operating department simulation, therefore not involving surgery. The study measured the contamination of settle plates as a method of recording bacterial dispersal.
Moore 2001	This study investigated the use of visors against masks. There were no surgical episodes where the surgical team's faces were uncovered. The surgical site infection rate was calculated on the outcome of a patient questionnaire. The subjective nature of these results meant that the study could not be used in the review.
Norman 1995	The use of visors and masks by staff was compared for acceptability and contamination. A group not wearing either mask or visor was not included.
Orr 1981	Excluded as it was not possible to distinguish how many clean operations were included in the study. Contact attempted with author.
Ritter 1975	This study was concerned with contamination of the environment rather than surgical site infection. Settle plates were used during non-operating period.
Ruthman 1984	The study examined the use of a cap and a mask in an Accident and Emergency department. These 2 variables could not be differentiated.
Salassa 2014	The study is not a randomised controlled trial; it is a review.
Sjol 2002	Stated as a RCT, but this study was observational and followed up patients for surgical wound infections post-discharge via a questionnaire.
Tunevall 1992	This study took place during actual operations but the specific outcome measure of the study was contamination of settle plates. Although it was reported that no surgical site infections occurred during the study period, the cross-over design of the study meant that all patients were exposed to a masked and non-masked period. The authors therefore could not utilise the results of this study.

ASA classification: the American Society of Anaesthesiologists physical status classification system is a system for assessing the fitness of patients before surgery
RCT: randomised controlled trial
SSI: surgical site infection

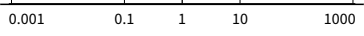
DATA AND ANALYSES

Comparison 1. Masks versus no masks

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Wound infection	3		Odds Ratio (M-H, Fixed, 95% CI)	Totals not selected

Analysis 1.1. Comparison 1 Masks versus no masks, Outcome 1 Wound infection.

Study or subgroup	Mask n/N	No mask n/N	Odds Ratio M-H, Fixed, 95% CI	Odds Ratio M-H, Fixed, 95% CI
Chamberlain 1984	0/14	3/10		0.07[0,1.63]
Tunevall 1991	13/706	10/723		1.34[0.58,3.07]
Webster 2010	33/313	31/340		1.17[0.7,1.97]



APPENDICES

Appendix 1. Search strategies

The Cochrane Wounds Specialised Register

#1 (mask or masks or facemask or facemasks or "face mask" or "face masks") AND (INREGISTER)
#2 (surg* NEAR5 (infect* or wound* or site* or incision* or dehisc*)) AND (INREGISTER)
#3 (wound* NEAR5 (infect* or site* or dehisc* or disrupt)) AND (INREGISTER)
#4 (wound NEXT complication*) AND (INREGISTER)
#5 #2 OR #3 OR #4
#6 #1 AND #5

The Cochrane Central Register of Controlled Trials (CENTRAL)

#1 MeSH descriptor: [Masks] explode all trees
#2 ("mask" or "masks" or facemask or facemasks or "face mask" or "face masks"):ti,ab,kw
#3 #1 or #2
#4 MeSH descriptor: [Surgical Wound Infection] explode all trees
#5 MeSH descriptor: [Surgical Wound Dehiscence] explode all trees
#6 (surg* near/5 infection*):ti,ab,kw
#7 (surg* near/5 wound*):ti,ab,kw
#8 (surg* near/5 site*):ti,ab,kw
#9 (surg* near/5 incision*):ti,ab,kw
#10 (surg* near/5 dehisc*):ti,ab,kw
#11 (wound* near/5 dehisc*):ti,ab,kw
#12 (wound* near/5 infect*):ti,ab,kw
#13 (wound near/5 disruption*):ti,ab,kw
#14 (wound next complication*):ti,ab,kw

#15 {or #4-#14}
#16 #3 and #15 in Trials

Ovid MEDLINE

1 exp Masks/
2 (mask*1 or facemask or face mask*).tw.
3 or/1-2
4 exp Surgical Wound Infection/
5 exp Surgical Wound Dehiscence/
6 (surg* adj5 infect*).tw.
7 (surg* adj5 wound*).tw.
8 (surg* adj5 site*).tw.
9 (surg* adj5 incision*).tw.
10 (surg* adj5 dehisc*).tw.
11 (wound* adj5 dehisc*).tw.
12 (wound* adj5 infect*).tw.
13 (wound adj5 disrupt*).tw.
14 wound complication*.tw.
15 or/4-14
16 3 and 15
17 randomized controlled trial.pt.
18 controlled clinical trial.pt.
19 randomi?ed.ab.
20 placebo.ab.
21 clinical trials as topic.sh.
22 randomly.ab.
23 trial.ti.
24 or/17-23
25 exp animals/ not humans.sh.
26 24 not 25
27 16 and 26

Ovid EMBASE

1 exp face mask/
2 (mask*1 or facemask or face mask*).tw.
3 or/1-2
4 exp surgical infection/
5 exp wound dehiscence/
6 (surg* adj5 infect*).tw.
7 (surg* adj5 wound*).tw.
8 (surg* adj5 site*).tw.
9 (surg* adj5 incision*).tw.
10 (surg* adj5 dehisc*).tw.
11 (wound* adj5 dehisc*).tw.
12 (wound* adj5 infect*).tw.
13 (wound adj5 disrupt*).tw.
14 wound complication*.tw.
15 or/4-14
16 3 and 15
17 Randomized controlled trials/
18 Single-Blind Method/
19 Double-Blind Method/
20 Crossover Procedure/
21 (random$ or factorial$ or crossover$ or cross over$ or cross-over$ or placebo$ or assign$ or allocat$ or volunteer$).ti,ab.
22 (doubl$ adj blind$).ti,ab.
23 (singl$ adj blind$).ti,ab.
24 or/17-23
25 exp animals/ or exp invertebrate/ or animal experiment/ or animal model/ or animal tissue/ or animal cell/ or nonhuman/
26 human/ or human cell/
27 and/25-26
28 25 not 27

29 24 not 28
30 16 and 29

EBSCO CINAHL Plus

S29 S16 AND S28
S28 S17 or S18 or S19 or S20 or S21 or S22 or S23 or S24 or S25 or S26 or S27
S27 MH "Quantitative Studies"
S26 TI placebo* or AB placebo*
S25 MH "Placebos"
S24TI random* allocat* or AB random* allocat*
S23 MH "Random Assignment"
S22 TI randomi?ed control* trial* or AB randomi?ed control* trial*
S21 AB (singl* or doubl* or trebl* or tripl*) and AB (blind* or mask*)
S20 TI (singl* or doubl* or trebl* or tripl*) and TI (blind* or mask*)
S19 TI clinic* N1 trial* or AB clinic* N1 trial*
S18 PT Clinical trial
S17 MH "Clinical Trials+"
S16 S3 AND S15
S15 S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14
S14 TI wound complication* or AB wound complication*
S13 TI wound* N5 disrupt* or AB wound* N5 disrupt*
S12 TI wound* N5 infect* or AB wound* N5 infect*
S11 TI wound* N5 dehisc* or AB wound* N5 dehisc*
S10 TI surg* N5 dehisc* or AB surg* N5 dehisc*
S9 TI surg* N5 incision* or AB surg* N5 incision*
S8 TI surg* N5 site* or AB surg* N5 site*
S7 TI surg* N5 wound* or AB surg* N5 wound*
S6 TI surg* N5 infect* or AB surg* N5 infect*
S5 (MH "Surgical Wound Dehiscence")
S4 (MH "Surgical Wound Infection")
S3 S1 or S2
S2 TI (mask* or facemask* or face mask) or AB (mask* or facemask* or face mask*)
S1 (MH "Masks")

Appendix 2. Risk of bias definitions

1. Was the allocation sequence randomly generated?

Low risk of bias

The investigators describe a random component in the sequence generation process such as: referring to a random number table; using a computer random number generator; coin tossing; shuffling cards or envelopes; throwing dice; drawing of lots.

High risk of bias

The investigators describe a non-random component in the sequence generation process. Usually, the description would involve some systematic, non-random approach, for example: sequence generated by odd or even date of birth; sequence generated by some rule based on date (or day) of admission; sequence generated by some rule based on hospital or clinic record number.

Unclear

Insufficient information about the sequence generation process to permit judgement of low or high risk of bias.

2. Was the treatment allocation adequately concealed?

Low risk of bias

Participants and investigators enrolling participants could not foresee assignment because one of the following, or an equivalent method, was used to conceal allocation: central allocation (including telephone, web-based and pharmacy-controlled randomisation); sequentially numbered drug containers of identical appearance; sequentially numbered, opaque, sealed envelopes.

High risk of bias

Participants or investigators enrolling participants could possibly foresee assignments and thus introduce selection bias, such as allocation based on: using an open random allocation schedule (e.g. a list of random numbers); assignment envelopes were used without appropriate

safeguards (e.g. if envelopes were unsealed or non-opaque or not sequentially numbered); alternation or rotation; date of birth; case record number; any other explicitly unconcealed procedure.

Unclear

Insufficient information to permit judgement of low or high risk of bias. This is usually the case if the method of concealment is not described or not described in sufficient detail to allow a definite judgement, for example if the use of assignment envelopes is described, but it remains unclear whether envelopes were sequentially numbered, opaque and sealed.

3. Blinding - was knowledge of the allocated interventions adequately prevented during the study?

Low risk of bias

Any one of the following.

- No blinding, but the review authors judge that the outcome and the outcome measurement are not likely to be influenced by lack of blinding.
- Blinding of participants and key study personnel ensured, and unlikely that the blinding could have been broken.
- Either participants or some key study personnel were not blinded, but outcome assessment was blinded and the non-blinding of others unlikely to introduce bias.

High risk of bias

Any one of the following.

- No blinding or incomplete blinding, and the outcome or outcome measurement is likely to be influenced by lack of blinding.
- Blinding of key study participants and personnel attempted, but likely that the blinding could have been broken.
- Either participants or some key study personnel were not blinded, and the non-blinding of others likely to introduce bias.

Unclear

Any one of the following.

- Insufficient information to permit judgement of low or high risk of bias.
- The study did not address this outcome.

4. Were incomplete outcome data adequately addressed?

Low risk of bias

Any one of the following.

- No missing outcome data.
- Reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to be introducing bias).
- Missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups.
- For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk not enough to have a clinically relevant impact on the intervention effect estimate.
- For continuous outcome data, plausible effect size (difference in means or standardised difference in means) among missing outcomes not enough to have a clinically relevant impact on observed effect size.
- Missing data have been imputed using appropriate methods.

High risk of bias

Any one of the following.

- Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups.
- For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce clinically relevant bias in intervention effect estimate.
- For continuous outcome data, plausible effect size (difference in means or standardised difference in means) among missing outcomes enough to induce clinically relevant bias in observed effect size.
- 'As-treated' analysis done with substantial departure of the intervention received from that assigned at randomisation.
- Potentially inappropriate application of simple imputation.

Unclear

Any one of the following.

- Insufficient reporting of attrition/exclusions to permit judgement of low or high risk of bias (e.g. number randomised not stated, no reasons for missing data provided).
- The study did not address this outcome.

5. Are reports of the study free of suggestion of selective outcome reporting?

Low risk of bias

Any of the following.

- The study protocol is available and all of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way.
- The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified (convincing text of this nature may be uncommon)

High risk of bias

Any one of the following.

- Not all of the study's pre-specified primary outcomes have been reported.
- One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified.
- One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect).
- One or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis.
- The study report fails to include results for a key outcome that would be expected to have been reported for such a study.

Unclear

Insufficient information to permit judgement of low or high risk of bias. It is likely that the majority of studies will fall into this category.

6. Other sources of potential bias

Low risk of bias

The study appears to be free of other sources of bias.

High risk of bias

There is at least one important risk of bias. For example, the study:

- had a potential source of bias related to the specific study design used; or
- had extreme baseline imbalance; or
- has been claimed to have been fraudulent; or
- had some other problem.

Unclear

There may be a risk of bias, but there is either:

- insufficient information to assess whether an important risk of bias exists; or
- insufficient rationale or evidence that an identified problem will introduce bias.

WHAT'S NEW

Date	Event	Description
1 April 2016	New citation required but conclusions have not changed	Seventh update; no change to conclusions; no new studies added.

Date	Event	Description
1 April 2016	New search has been performed	New search.

HISTORY

Protocol first published: Issue 4, 2000
Review first published: Issue 1, 2002

Date	Event	Description
29 October 2013	New search has been performed	Sixth update.
29 October 2013	New citation required but conclusions have not changed	New search; no new studies identified; no change to conclusions.
19 January 2010	New search has been performed	New search; one additional trial included (Webster 2010); no change to conclusions. Clarification of participants being the patients undergoing surgery not the members of the surgical team wearing the face mask.
18 June 2008	Amended	Converted to new review format.
4 February 2008	New search has been performed	For this third update new searches were carried out in February 2008. No new relevant studies were identified. The authors' conclusions remain unchanged. Published in *The Cochrane Library*, Issue 2, 2008.
10 February 2006	New search has been performed	For the second update new searches were carried out in February 2006. One new study was identified (Alwitry 2002), but was excluded from the review. Published in *The Cochrane Library*, Issue 3, 2006.
16 April 2004	New search has been performed	For the first update, new searches were carried out in April 2004. One new study was identified (Sjol 2002), but was excluded from the review. Published in *The Cochrane Library*, Issue 3, 2004.
20 November 2001	New citation required and conclusions have changed	Substantive amendment.

CONTRIBUTIONS OF AUTHORS

Peggy Edwards identified studies from the initial search and selected studies independently for data extraction, devised the data extraction sheet, independently extracted the data from studies, drafted the protocol and the review jointly with Allyson Lipp, provided content expertise and agreed with the update of the review.

Marina Vincent undertook the seventh update of this review, screened the search output and updated the text and plain language summary.

Contributions of editorial base

Nicky Cullum: edited the review, advised on methodology, interpretation and review content. Approved the final review and review update prior to submission.
Sally Bell-Syer: co-ordinated the editorial process. Advised on methodology, interpretation and content. Edited the review and the updated review.
Ruth Foxlee: designed the search strategy, ran the searches and edited the search methods section for previous updates.
Rachel Richardson: checked previous review updates prior to submission.

Reetu Child: ran the searches and checked the search strategy for this update.

DECLARATIONS OF INTEREST

Marina Vincent: none known.
Peggy Edwards: none known.

SOURCES OF SUPPORT

Internal sources

- Faculty of Health, Sport and Science University of Glamorgan, UK.

External sources

- Theatre Nursing Trust Fund, UK.
- National Institute for Health Research (NIHR), UK.
- This project was supported by the National Institute for Health Research via Cochrane Infrastructure funding to Cochrane Wounds. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, NHS or the Department of Health, UK.

INDEX TERMS

Medical Subject Headings (MeSH)

*Masks; Disposable Equipment; Randomized Controlled Trials as Topic; Surgical Wound Infection [*prevention & control]

MeSH check words

Humans

Infect Control Hosp Epidemiol. 1997 Jan;18(1):49-57. doi: 10.2307/30141964.

The evolution of the surgical mask: filtering efficiency versus effectiveness

N L Belkin

PMID: 9013247 DOI: 10.2307/30141964

Abstract

When originally introduced for use at the turn of the century, the primary function of the surgical mask was to prevent the migration of microorganisms residing in the nose and mouth of members of the operating team to the open wound of the patient. As technology developed new materials and designs, their filtering efficiencies gradually improved. However, there is no standard test method for assessing that capability, and its influence on the rates of surgical-wound infection has yet to be demonstrated. Quite to the contrary, both in-vitro and in-vivo studies indicate that a mask may not be universally necessary in today's surgical environment.

Comment in

Making surgical care better: hard work, small gains.
Lee JT. Lee JT. Infect Control Hosp Epidemiol. 1997 Jan;18(1):6-8. doi: 10.1086/647493. Infect Control Hosp Epidemiol. 1997. PMID: 9013239 No abstract available.

Related information

MedGen

LinkOut - more resources

Medical
MedlinePlus Health Information

SHORT REPORT | VOLUME 18, ISSUE 3, P239-242, JULY 01, 1991

Surgical face masks in modern operating rooms—a costly and unnecessary ritual?

N.J. Mitchell ● S. Hunt DOI: https://doi.org/10.1016/0195-6701(91)90148-2

Abstract

Following the commissioning of a new suite of operating rooms air movement studies showed a flow of air away from the operating table towards the periphery of the room. Oral microbial flora dispersed by unmasked male and female volunteers standing one metre from the table failed to contaminate exposed settle plates placed on the table. The wearing of face masks by non-scrubbed staff working in an operating room with forced ventilation seems to be unnecessary.

Keywords

Face masks • operating theatre • bacterial dispersal

References

1. Rogers KB
An investigation into the efficacy of disposable face masks.
J Clin Pathol. 1980; **33**: 1086-1091

Scopus (21) • PubMed • Crossref • Google Scholar

2. Shah M • Crompton P • Vickers MDA
The efficacy of face masks.
Ann Roy Coll Surg Engl. 1983; **65**: 380-381

PubMed • Google Scholar

3. Ha'evi GB • Wiley AM
e efficacy of standard surgical face masks: an investigation using "tracer cle "
n Orthop. 1980; **148**: 160-162

PubMed • Google Scholar

4. Ritter MA • Eitzen H • French MLV • Hart JB
The operating room environment as affected by people and the surgical face mask.
Clin Orthop. 1975; **111**: 147-150

Scopus (93) • PubMed • Crossref • Google Scholar

5. Hare R • Thomas CGA
The transmission of *Staphylococcus aureus*.
Br Med J. 1956; **232**: 840-844

Scopus (50) • Crossref • Google Scholar

6. Shooter RA • Smith MA • Hunter CJW
A study of surgical masks.
Br J Surg. 1959; **47**: 246-249

Scopus (17) • PubMed • Crossref • Google Scholar

7. Orr NWM
Is a mask necessary in the operating theatre?.
Ann Roy Coll Surg Engl. 1981; **63**: 380-392

Google Scholar

8. Ruthman JC • Hendricksen D • Miller RF • Quigg DL
Effect of cap and mask on infection rates.
Illinois Medical Journal. 1984; **165**: 397-399

Google Scholar

9. Laslett LJ • Sabin A
Wearing of caps and masks not necessary during cardiac catheterization.
ıthet Cardiovasc Diagn. 1989; **17**: 158-160

Article Info

Publication History

Accepted: May 22, 1991

Identification

DOI: https://doi.org/10.1016/0195-6701(91)90148-2

Copyright

ScienceDirect

Related Articles

Effectiveness of surgical masks against influenza bioaerosols

Journal of Hospital Infection, Vol. 84, Issue 1

In Brief • Full-Text • PDF

The use of germicidal ultraviolet light, vaporised hydrogen peroxide and dry heat to decontaminate face masks and filtering respirators contaminated with a SARS-CoV-2 surrogate virus

Journal of Hospital Infection

In Brief • Full-Text • PDF

Effect of surgical mask position on bacterial contamination of the operative field

Journal of Hospital Infection, Vol. 23, Issue 1

In Brief • Full-Text • PDF

Use of ultraviolet-fluorescence-based simulation in evaluation of personal protective equipment worn for first assessment and care of a patient with suspected high-consequence infectious disease

Journal of Hospital Infection, Vol. 99, Issue 2

In Brief • Full-Text • PDF

Open Access

Dry heat and microwave-generated steam protocols for the rapid decontamination of respiratory personal protective equipment in response to COVID-19-related shortages

Journal of Hospital Infection, Vol. 106, Issue 1

In Brief • Full-Text • PDF

Antimicrobial effect of surgical masks coated with nanoparticles

Journal of Hospital Infection, Vol. 62, Issue 1

In Brief • Full-Text • PDF

What face mask for what use in the context of the COVID-19 pandemic? The French guidelines

Journal of Hospital Infection, Vol. 105, Issue 3

f • Full-Text • PDF

Cathet Cardiovasc Diagn. 1989 Jul;17(3):158-60. doi: 10.1002/ccd.1810170306.

Wearing of caps and masks not necessary during cardiac catheterization

L J Laslett [1], A Sabin

Affiliations
PMID: 2766345 DOI: 10.1002/ccd.1810170306

Abstract

Although cardiac catheterization-related infections are rare, caps and masks are often worn to minimize this complication. However, documentation of the value of caps and masks for this purpose is lacking. We, therefore, prospectively evaluated the experience of 504 patients undergoing percutaneous left heart catheterization, seeking evidence of a relationship between whether caps and/or masks were worn by the operators and the incidence of infection. No infections were found in any patient, regardless of whether a cap or mask was used. Thus, we found no evidence that caps or masks need to be worn during percutaneous cardiac catheterization.

Related information

MedGen

LinkOut - more resources

Full Text Sources

Wiley

Miscellaneous

NCI CPTAC Assay Portal

COMMENTARY: Masks-for-all for COVID-19 not based on sound data

Filed Under: **COVID-19** (/infectious-disease-topics/covid-19)

Lisa M Brosseau, ScD, and Margaret Sietsema, PhD (/ongoing-programs/news-publishing/news-publishing-staff) | Apr 01, 2020

Dr. Brosseau (https://www.cidrap.umn.edu/about-us/staff/lisa-m-brosseau-scd-cih) is a national expert on respiratory protection and infectious diseases and professor (retired), University of Illinois at Chicago.
Dr. Sietsema (https://publichealth.uic.edu/?s=sietsema) is also an expert on respiratory protection and an assistant professor at the University of Illinois at Chicago.

Vergani_Fotografia / iStoc

Editor's Note: *The authors added the following statement on Jul 16.*

The authors and CIDRAP have received requests in recent weeks to remove this article from the CIDRAP website. Reasons have included: (1) we don't truly know that cloth masks (face coverings) are not effective, since the data are so limited, (2) wearing a cloth mask or face covering is better than doing nothing, (3) the article is being used by individuals and groups to support non-mask wearing where mandated and (4) there are now many modeling studies suggesting that cloth masks or face coverings could be effective at flattening the curve and preventing many cases of infection.

If the data are limited, how can we say face coverings are likely not effective?

We agree that the data supporting the effectiveness of a cloth mask or face covering are very limited. We do, however, have data from laboratory studies that indicate cloth masks or face coverings offer very low filter collection efficiency for the smaller inhalable particles we believe are largely responsible for transmission, particularly from pre- or asymptomatic individuals who are not coughing or sneezing. At the time we wrote this article, we were unable to locate any well-performed studies of cloth mask leakage when worn on the face—either inward or outward leakage. As far as we know, these data are still lacking.

The guidelines from the Centers for Disease Control and Prevention (CDC) for face coverings initially did not have any citations for studies of cloth material efficiency or fit, but some references have been added since the guidelines were first posted. We reviewed these and found that many employ very crude, non-standardized methods (Anfinrud 2020, Davies 2013, Konda 2020, Aydin 2020, Ma 2020) or are not relevant to cloth face coverings because they evaluate respirators or surgical masks (Leung 2020, Johnson 2009, Green 2012).

The CDC failed to reference the National Academies of Sciences Rapid Expert Consultation on the Effectiveness of Fabric Masks for the COVID-19 Pandemic (NAS 2020), which concludes, "The

evidence from…laboratory filtration studies suggests that such fabric masks may reduce the transmission of larger respiratory droplets. There is little evidence regarding the transmission of small aerosolized particulates of the size potentially exhaled by asymptomatic or presymptomatic individuals with COVID-19." As well, the CDC neglected to mention a well-done study of cloth material filter performance by Rengasamy et al (2014), which we reviewed in our article.

Is wearing a face covering better than nothing?

Wearing a cloth mask or face covering could be better than doing nothing, but we simply don't know at this point. We have observed an evolution in the messaging around cloth masks, from an initial understanding that they should not be seen as a replacement for physical distancing to more recent messaging that suggests cloth masks are equivalent to physical distancing. And while everyone appears to understand that this messaging suggests that a cloth mask is appropriate only for source control (ie, to protect others from infection), recent CDC and other guidance recommending their use by workers seems to imply that they offer some type of personal protection.

We know of workplaces in which employees are told they cannot wear respirators for the hazardous environments they work in, but instead need to wear a cloth mask or face covering. These are dangerous and inappropriate applications that greatly exceed the initial purpose of a cloth mask. We are concerned that many people do not understand the very limited degree of protection a cloth mask or face covering likely offers as source control for people located nearby.

Do we support cloth mask wearing where mandated?

Despite the current limited scientific data detailing their effectiveness, we support the wearing of face coverings by the public when mandated and when in close contact with people whose infection status they don't know. However, we also encourage everyone to continue to limit their time spent indoors near potentially infectious people and to not count on or expect a cloth mask or face covering to protect them or the people around them. The pandemic is not over and will not likely be over for some time. As states and local jurisdictions reopen, we encourage people to continue to assess and limit their risks. Cloth masks and face coverings likely do *not* offer the same degree of protection as physical distancing, isolation, or limiting personal contact time.

Will face coverings 'flatten the curve' and stop the pandemic?

We have reviewed the many modeling studies that purport to demonstrate that cloth masks or face coverings have the potential for flattening the curve or significantly decrease the number of cases. These studies fail to recognize several important facts:

- The filter performance of a cloth material does not directly translate or represent its performance on an individual, because it neglects the understanding of fit.
- Cloth masks or coverings come in a variety of shapes, sizes, and materials and are not made according to any standards.
- Transmission is not simply a function of short random interactions between individuals, but rather a function of particle concentration in the air and the time exposed to that concentration.
- A cloth mask or face covering does very little to prevent the emission or inhalation of small particles. As discussed in an earlier CIDRAP commentary (https://www.cidrap.umn.edu/news-perspective/2020/03/commentary-covid-19-transmission-messages-should-hinge-science) and more recently by Morawska and Milton (2020) in an open letter to WHO signed by 239 scientists, inhalation of

small infectious particles is not only biologically plausible, but the epidemiology supports it as an important mode of transmission for SARS-CoV-2, the virus that causes COVID-19.

In summary, though we support mask wearing by the general public, we continue to conclude that cloth masks and face coverings are likely to have limited impact on lowering COVID-19 transmission, because they have minimal ability to prevent the emission of small particles, offer limited personal protection with respect to small particle inhalation, and should not be recommended as a replacement for physical distancing or reducing time in enclosed spaces with many potentially infectious people. We are very concerned about messaging that suggests cloth masks or face coverings can replace physical distancing. We also worry that the public doesn't understand the limitations of cloth masks and face coverings when we observe how many people wear their mask under their nose or even under their mouth, remove their masks when talking to someone nearby, or fail to practice physical distancing when wearing a mask.

References

Anfinrud P, Stadnytskyi V, Bax CE, et al. Visualizing speech-generated oral fluid droplets with laser light scattering (https://www.nejm.org/doi/full/10.1056/nejmc2007800) . N Engl J Med 2020 (published online Apr 15)

Davies A, Thompson KA, Giri K, et al. Testing the efficacy of homemade masks: would they protect in an influenza pandemic? (https://www.cambridge.org/core/journals/disaster-medicine-and-public-health-preparedness/article/testing-the-efficacy-of-homemade-masks-would-they-protect-in-an-influenza-pandemic/0921A05A69A9419C862FA2F35F819D55) Disaster Med Public Health Prep 2013 Aug;7(4):413-8

Green CF, Davidson CS, Panlilio AL, et al. Effectiveness of selected surgical masks in arresting vegetative cells and endospores when worn by simulated contagious patients (https://www.cambridge.org/core/journals/infection-control-and-hospital-epidemiology/article/effectiveness-of-selected-surgical-masks-in-arresting-vegetative-cells-and-endospores-when-worn-by-simulated-contagious-patients/BFC344C8AF4D992EB843B1FF0644BC84) . Infect Control Hosp Epidemiol 2012 May;33(5):487-94

Johnson DF, Druce JD, Birch C, et al. A quantitative assessment of the efficacy of surgical and N95 masks to filter influenza virus in patients with acute influenza infection (https://academic.oup.com/cid/article/49/2/275/405108) . Clin Infect Dis 2009 Jul 15;49(2):275-7

Konda A, Prakash A, Moss GA, et al. Aerosol filtration efficiency of common fabrics used in respiratory cloth masks (https://pubs.acs.org/doi/10.1021/acsnano.0c03252) . ACS Nano. 2020 (published online Apr 24)

Leung NHL, Chu DKW, Shiu EYC, et al. Respiratory virus shedding in exhaled breath and efficacy of face masks (https://www.nature.com/articles/s41591-020-0843-2) . *Nat Med* 2020 (published online Apr 3)

Ma QX, Shan H, Zhang HL, et al. Potential utilities of mask-wearing and instant hand hygiene for fighting SARS-CoV-2 (https://onlinelibrary.wiley.com/doi/full/10.1002/jmv.25805) . J Med Virol 2020 (published online Mar 31)

Morawska L, Milton DK. It is time to address airborne transmission of COVID-19. Clin Infect Dis 2020 (published online Jul 6)

National Academies of Sciences, Engineering, and Medicine. 2020. Rapid expert consultation on the effectiveness of fabric masks for the COVID-19 pandemic (https://www.nap.edu/catalog/25776/rapid-expert-consultation-on-the-effectiveness-of-fabric-masks-for-the-covid-19-pandemic-april-8-2020) . Washington, DC, National Academies Press. Apr 8, 2020

Rengasamy S, Eimer B, Szalajda J. A quantitative assessment of the total inward leakage of NaCl aerosol representing submicron-size bioaerosol through N95 filtering facepiece respirators and surgical masks (https://www.tandfonline.com/doi/full/10.1080/15459624.2013.866715) . J Occup Environ Hyg 2014 May 9;11(6):388-96

Editor's Note: *Also on Jul 16, The following text was changed directly after the "Surgical masks as source control" subhead in the original commentary:*

Original: *Household studies find very limited effectiveness of surgical masks at reducing respiratory illness in other household members.*[22-25]

Updated: *We were able to identify only two household studies in which surgical masks were worn by the index patient only, as source control.*[24,25] *Neither of these found a significant impact on secondary disease transmission, although both studies had important limitations.*

The original reference 24 (bin-Reza 2011) was changed to Canini 2010. In an unrelated correction on Jul 16, reference 45 was incorrect and now correctly cites bin-Reza 2012.

In response to the stream of misinformation and misunderstanding about the nature and role of masks and respirators as source control or personal protective equipment (PPE), we critically review the topic to inform ongoing COVID-19 decision-making that relies on science-based data and professional expertise.

As noted in a previous commentary (http://www.cidrap.umn.edu/news-perspective/2020/03/commentary-covid-19-transmission-messages-should-hinge-science) , the limited data we have for COVID-19 strongly support the possibility that SARS-CoV-2—the virus that causes COVID-19—is transmitted by inhalation of both droplets and aerosols near the source. It is also likely that people who are pre-symptomatic or asymptomatic throughout the duration of their infection are spreading the disease in this way.

Data lacking to recommend broad mask use

We do not recommend requiring the general public who do not have symptoms of COVID-19-like illness to routinely wear cloth or surgical masks because:

- There is no scientific evidence they are effective in reducing the risk of SARS-CoV-2 transmission
- Their use may result in those wearing the masks to relax other distancing efforts because they have a sense of protection
- We need to preserve the supply of surgical masks for at-risk healthcare workers.

Sweeping mask recommendations—as many have proposed—will not reduce SARS-CoV-2 transmission, as evidenced by the widespread practice of wearing such masks in Hubei province, China, before and during its mass COVID-19 transmission experience earlier this year. Our review of relevant studies indicates that cloth masks will be ineffective at preventing SARS-CoV-2 transmission, whether worn as source control or as PPE.

Surgical masks likely have some utility as source control (meaning the wearer limits virus dispersal to another person) from a symptomatic patient in a healthcare setting to stop the spread of large cough particles and limit the lateral dispersion of cough particles. They may also have very limited utility as source control or PPE in households.

Respirators, though, are the only option that can ensure protection for frontline workers dealing with COVID-19 cases, once all of the strategies (https://www.cdc.gov/coronavirus/2019-ncov/hcp/respirators-strategy/index.html) for optimizing respirator supply have been implemented.

We do not know whether respirators are an effective intervention as source control for the public. A non-fit-tested respirator may not offer any better protection than a surgical mask. Respirators work as PPE only when they are the right size and have been fit-tested to demonstrate they achieve an adequate protection factor. In a time when respirator supplies are limited, we should be saving them for frontline workers to prevent infection and remain in their jobs.

These recommendations are based on a review of available literature and informed by professional expertise and consultation. We outline our review criteria, summarize the literature that best addresses these criteria, and describe some activities the public can do to help "flatten the curve" and to protect frontline workers and the general public.

We realize that the public yearns to help protect medical professionals by contributing homemade masks, but there are better ways to help.

Filter efficiency and fit are key for masks, respirators

The best evidence of mask and respirator performance starts with testing filter efficiency and then evaluating fit (facepiece leakage). Filter efficiency must be measured first. If the filter is inefficient, then fit will be a measure of filter efficiency only and not what is being leaked around the facepiece.

Filter efficiency

Masks and respirators work by collecting particles through several physical mechanisms, including diffusion (small particles) and interception and impaction (large particles).[1] N95 filtering facepiece respirators (FFRs) are constructed from electret filter material, with electrostatic attraction for additional collection of all particle sizes.[2]

Every filter has a particle size range that it collects inefficiently. *Above and below this range, particles will be collected with greater efficiency*. For fibrous non-electret filters, this size is about 0.3 micrometers (µm); for electret filters, it ranges from 0.06 to 0.1 µm. When testing, we care most about the point of inefficiency. As flow increases, particles in this range will be collected less efficiently.

The best filter tests use worst-case conditions: high flow rates (80 to 90 liters per minute [L/min]) with particle sizes in the least efficiency range. This guarantees that filter efficiency will be high at

typical, lower flow rates for all particle sizes. Respirator filter certification tests use 84 L/min, well above the typical 10 to 30 L/min breathing rates. The N95 designation means the filter exhibits at least 95% efficiency in the least efficient particle size range.

Studies should also use well-characterized inert particles (not biological, anthropogenic, or naturogenic ones) and instruments that quantify concentrations in narrow size categories, and they should include an N95 FFR or similar respirator as a positive control.

Fit

Fit should be a measure of how well the mask or respirator prevents leakage around the facepiece, as noted earlier. Panels of representative human subjects reveal more about fit than tests on a few individuals or mannequins.

Quantitative fit tests that measure concentrations inside and outside of the facepiece are more discriminating than qualitative ones that rely on taste or odor.

Mask, N95 respirator filtering performance

Following a recommendation that cloth masks be explored for use in healthcare settings during the next influenza pandemic,[3] The National Institute for Occupational Safety and Health (NIOSH) conducted a study of the filter performance on clothing materials and articles, including commercial cloth masks marketed for air pollution and allergens, sweatshirts, t-shirts, and scarfs.[4]

Filter efficiency was measured across a wide range of small particle sizes (0.02 to 1 µm) at 33 and 99 L/min. N95 respirators had efficiencies greater than 95% (as expected). For the entire range of particles tested, t-shirts had 10% efficiency, scarves 10% to 20%, cloth masks 10% to 30%, sweatshirts 20% to 40%, and towels 40%. All of the cloth masks and materials had near zero efficiency at 0.3 µm, a particle size that easily penetrates into the lungs.[4]

Another study evaluated 44 masks, respirators, and other materials with similar methods and small aerosols (0.08 and 0.22 µm).[5] N95 FFR filter efficiency was greater than 95%. Medical masks exhibited 55% efficiency, general masks 38% and handkerchiefs 2% (one layer) to 13% (four layers).

These studies demonstrate that cloth or homemade masks will have very low filter efficiency (2% to 38%). Medical masks are made from a wide range of materials, and studies have found a wide range of filter efficiency (2% to 98%), with most exhibiting 30% to 50% efficiency.[6-12]

We reviewed other filter efficiency studies of makeshift cloth masks made with various materials. Limitations included challenge aerosols that were poorly characterized[13] or too large[14-16] or flow rates that were too low.[17]

Mask and respirator fit

Regulators have not developed guidelines for cloth or surgical mask fit. N95 FFRs must achieve a fit factor (outside divided by inside concentration) of at least 100, which means that the facepiece must lower the outside concentration by 99%, according to the OSHA respiratory protection standard (https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.134). When fit is measured on a mask with inefficient filters, it is really a measure of the collection of particles by the filter plus how well the mask prevents particles from leaking around the facepiece.

Several studies have measured the fit of masks made of cloth and other homemade materials.[13,18,19] We have not used their results to evaluate mask performance, because none measured filter efficiency or included respirators as positive controls.

One study of surgical masks showing relatively high efficiencies of 70% to 95% using NIOSH test methods measured total mask efficiencies (filter plus facepiece) of 67% to 90%.[7] These results illustrate that surgical masks, even with relatively efficient filters, do not fit well against the face.

In sum, cloth masks exhibit very low filter efficiency. Thus, even masks that fit well against the face will not prevent inhalation of small particles by the wearer or emission of small particles from the wearer.

One study of surgical mask fit described above suggests that poor fit can be somewhat offset by good filter collection, but will not approach the level of protection offered by a respirator. The problem is, however, that many surgical masks have very poor filter performance. Surgical masks are not evaluated using worst-case filter tests, so there is no way to know which ones offer better filter efficiency.

Studies of performance in real-world settings

Before recommending them, it's important to understand how masks and respirators perform in households, healthcare, and other settings.

Cloth masks as source control

A historical overview of cloth masks notes their use in US healthcare settings starting in the late 1800s, first as source control on patients and nurses and later as PPE by nurses.[20]

Kellogg,[21] seeking a reason for the failure of cloth masks required for the public in stopping the 1918 influenza pandemic, found that the number of cloth layers needed to achieve acceptable efficiency made them difficult to breathe through and caused leakage around the mask. We found no well-designed studies of cloth masks as source control in household or healthcare settings.

In sum, given the paucity of information about their performance as source control in real-world settings, along with the extremely low efficiency of cloth masks as filters and their poor fit, there is no evidence to support their use by the public or healthcare workers to control the emission of particles from the wearer.

Surgical masks as source control

We were able to identify only two household studies in which surgical masks were worn by the index patient only, as source control.[24,25] Neither of these found a significant impact on secondary disease transmission, although both studies had important limitations.

Clinical trials in the surgery theater have found no difference in wound infection rates with and without surgical masks.[26-29] Despite these findings, it has been difficult for surgeons to give up a long-standing practice.[30]

There is evidence from laboratory studies with coughing infectious subjects that surgical masks are effective at preventing emission of large particles[31-34] and minimizing lateral dispersion of cough

particles, but with simultaneous displacement of aerosol emission upward and downward from the mask.[35]

There is some evidence that surgical masks can be effective at reducing overall particle emission from patients who have multidrug-resistant tuberculosis,[36] cystic fibrosis,[34] and influenza.[33] The latter found surgical masks decreased emission of large particles (larger than 5 µm) by 25-fold and small particles by threefold from flu-infected patients.[33] Sung[37] found a 43% reduction in respiratory viral infections in stem-cell patients when everyone, including patients, visitors, and healthcare workers, wore surgical masks.

In sum, wearing surgical masks in households appears to have very little impact on transmission of respiratory disease. One possible reason may be that masks are not likely worn continuously in households. These data suggest that surgical masks worn by the public will have no or very low impact on disease transmission during a pandemic.

There is no evidence that surgical masks worn by healthcare workers are effective at limiting the emission of small particles or in preventing contamination of wounds during surgery.

There is moderate evidence that surgical masks worn by patients in healthcare settings can lower the emission of large particles generated during coughing and limited evidence that small particle emission may also be reduced.

N95 FFRs as source control

Respirator use by the public was reviewed by NIOSH (https://blogs.cdc.gov/niosh-science-blog/2018/01/04/respirators-public-use/) : (1) untrained users will not wear respirators correctly, (2) non-fit tested respirators are not likely to fit, and (3) improvised cloth masks do not provide the level of protection of a fit-tested respirator.

There are few studies examining the effectiveness of respirators on patients. An N95 FFR on coughing human subjects showed greater effectiveness at limiting lateral particle dispersion than surgical masks (15 cm and 30 cm dispersion, respectively) in comparison to no mask (68 cm). [35] Cystic fibrosis patients reported that surgical masks were tolerable for short periods, but N95 FFRs were not.[34]

In summary, N95 FFRs on patients will not be effective and may not be appropriate, particularly if they have respiratory illness or other underlying health conditions. Given the current extreme shortages of respirators needed in healthcare, we do not recommend the use of N95 FFRs in public or household settings.

Cloth masks as PPE

A randomized trial comparing the effect of medical and cloth masks on healthcare worker illness found that those wearing cloth masks were 13 times more likely to experience influenza-like illness than those wearing medical masks.[38]

In sum, very poor filter and fit performance of cloth masks described earlier and very low effectiveness for cloth masks in healthcare settings lead us conclude that cloth masks offer no protection for healthcare workers inhaling infectious particles near an infected or confirmed patient.

Surgical masks as PPE

Several randomized trials have not found any statistical difference in the efficacy of surgical masks versus N95 FFRs at lowering infectious respiratory disease outcomes for healthcare workers.[39-43]

Most reviews have failed to find any advantage of one intervention over the other.[23,44-48] Recent meta-analyses found that N95 FFRs offered higher protection against clinical respiratory illness[49,50] and lab-confirmed bacterial infections,[49] but not viral infections or influenza-like illness.[49]

A recent pooled analysis of two earlier trials comparing medical masks and N95 filtering facepiece respirators **with controls** (no protection) found that healthcare workers continuously wearing N95 FFRs were 54% less likely to experience respiratory viral infections than controls ($P = 0.03$), while those wearing medical masks were only 12% less likely than controls ($P = 0.48$; result is not significantly different from zero).[51]

While the data supporting the use of surgical masks as PPE in real-world settings are limited, the two meta-analyses and the most recent randomized controlled study[51] combined with evidence of moderate filter efficiency and complete lack of facepiece fit lead us to conclude that surgical masks offer very low levels of protection for the wearer from aerosol inhalation. There may be some protection from droplets and liquids propelled directly onto the mask, but a faceshield would be a better choice if this is a concern.

N95 FFRs as PPE

A retrospective cohort study found that nurses' risk of SARS (severe acute respiratory syndrome, also caused by a coronavirus) was lower with consistent use of N95 FFRs than with consistent use of a surgical mask.[52]

In sum, this study, the meta-analyses, randomized controlled trial described above,[49,51] and laboratory data showing high filter efficiency and high achievable fit factors lead us to conclude that N95 FFRs offer superior protection from inhalable infectious aerosols likely to be encountered when caring for suspected or confirmed COVID-19 patients.

The precautionary principle supports higher levels of respiratory protection, such as powered air-purifying respirators, for aerosol-generating procedures such as intubation, bronchoscopy, and acquiring respiratory specimens.

Conclusions

While this is not an exhaustive review of masks and respirators as source control and PPE, we made our best effort to locate and review the most relevant studies of laboratory and real-world performance to inform our recommendations. Results from laboratory studies of filter and fit performance inform and support the findings in real-world settings.

Cloth masks are ineffective as source control and PPE, surgical masks have some role to play in preventing emissions from infected patients, and respirators are the best choice for protecting healthcare and other frontline workers, but not recommended for source control. These recommendations apply to pandemic and non-pandemic situations.

Leaving aside the fact that they are ineffective, telling the public to wear cloth or surgical masks could be interpreted by some to mean that people are safe to stop isolating at home. It's too late now for anything but stopping as much person-to-person interaction as possible.

Masks may confuse that message and give people a false sense of security. If masks had been the solution in Asia, shouldn't they have stopped the pandemic before it spread elsewhere?

Ways to best protect health workers

We recommend that healthcare organizations follow US Centers for Disease Control and Prevention (CDC) guidance (https://www.cdc.gov/coronavirus/2019-ncov/hcp/respirators-strategy/index.html) by moving first through conventional, then contingency, and finally crisis scenarios to optimize the supply of respirators. We recommend using the CDC's burn rate calculator (https://www.cdc.gov/coronavirus/2019-ncov/hcp/ppe-strategy/burn-calculator.html) to help identify areas to reduce N95 consumption and working down the CDC checklist (https://www.cdc.gov/coronavirus/2019-ncov/hcp/checklist-n95-strategy.html) for a strategic approach to extend N95 supply.

For readers who are disappointed in our recommendations to stop making cloth masks for themselves or healthcare workers, we recommend instead pitching in to locate N95 FFRs and other types of respirators for healthcare organizations. Encourage your local or state government to organize and reach out to industries to locate respirators not currently being used in the non-healthcare sector and coordinate donation efforts (https://www.cnbc.com/2020/03/25/apple-and-facebook-face-masks-were-stockpiled-after-wildfires.html) to frontline health workers.

References

1. **Lee KW, Liu BYH.** On the minimum efficiency and the most penetrating particle size for fibrous filters (https://www.tandfonline.com/doi/abs/10.1080/00022470.1980.10464592) . J Air Pollut Control Assoc 1980 Mar 13;30(4):377-81
2. **Martin SB Jr, Moyer ES.** Electrostatic respirator filter media: filter efficiency and most penetrating particle size effects (https://www.tandfonline.com/doi/abs/10.1080/10473220050075617) . Appl Occup Environ Hyg 2000 Nov 30;15(8):609-17
3. Reusability of facemasks during an influenza pandemic. (https://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=s04272006) News conference, Apr 27, 2006
4. **Rengasamy S, Eimer B, Shaffer RE.** Simple respiratory protection—evaluation of the filtration performance of cloth masks and common fabric materials against 20-1000 nm size particles. (https://academic.oup.com/annweh/article/54/7/789/202744) Ann Occup Hyg 2010 Jun 28;54(7):789-98
5. **Jung H, Kim J, Lee S, et al.** Comparison of filtration efficiency and pressure drop in anti-yellow sand masks, quarantine masks, medical masks, general masks, and handkerchiefs. (http://aaqr.org/files/article/668/36_AAQR-13-06-OA-0201_991-1002.pdf) Aerosol Air Qual Res 2014;14(14):991-1002.
6. **Grinshpun SA, Haruta H, Eninger RM, et al.** Performance of an N95 filtering facepiece particulate respirator and a surgical mask during human breathing: two pathways for particle penetration (https://oeh.tandfonline.com/doi/full/10.1080/15459620903120086) . J Occup Environ Hyg 2009 Jul 22;6(10):593-603

7. **Oberg T, Brosseau LM.** Surgical mask filter and fit performance (https://www.ajicjournal.org/article/S0196-6553(07)00774-2/fulltext) . Am J Infect Control 2008 May;36(4):276-82
8. **Willeke K, Qian Y, Donnelly J, et al.** Penetration of airborne microorganisms through a surgical mask and a dust/mist respirator (https://www.tandfonline.com/doi/abs/10.1080/15428119691014882) . Am Ind Hyg Assoc J 1996;57(4):348-55
9. **Brosseau LM, McCullough NV, Vesley D.** Mycobacterial aerosol collection efficiency of respirator and surgical mask filters under varying conditions of flow and humidity (https://www.tandfonline.com/doi/abs/10.1080/1047322X.1997.10389533) . Appl Occup Environ Hyg 1997;12(6):435-45
10. **Chen CC, Willeke K.** Aerosol penetration through surgical masks (https://www.ajicjournal.org/article/S0196-6553(05)80143-9/fulltext) . Am J Infect Control 1992 Aug;20(4):177-84
11. **McCullough NV, Brosseau LM, Vesley D.** Collection of three bacterial aerosols by respirator and surgical mask filters under varying conditions of flow and relative humidity (https://www.sciencedirect.com/science/article/abs/pii/S0003487897000227) . Ann Occup Hyg 1997 Dec;41(6):677-90
12. **Rengasamy S, Eimer B, Szalajda J.** A quantitative assessment of the total inward leakage of NaCl aerosol representing submicron-size bioaerosol through N95 filtering facepiece respirators and surgical masks (https://www.tandfonline.com/doi/full/10.1080/15459624.2013.866715) . J Occup Environ Hyg 2014 11(6):388-96
13. **Davies A, Thompson KA, Giri K, et al.** Testing the efficacy of homemade masks: would they protect in an influenza pandemic? (https://www.cambridge.org/core/journals/disaster-medicine-and-public-health-preparedness/article/testing-the-efficacy-of-homemade-masks-would-they-protect-in-an-influenza-pandemic/0921A05A69A9419C862FA2F35F819D55) Disaster Med Public Health Prep 2013 Aug;7(4):413-8
14. Cherrie JW, Apsley A, Cowie H, et al. Effectiveness of face masks used to protect Beijing residents against particulate air pollution. (https://oem.bmj.com/content/75/6/446.long) Occup Environ Med 2018 Jun;75(6):446-52
15. **Mueller W, Horwell CJ, Apsley A, et al.** The effectiveness of respiratory protection worn by communities to protect from volcanic ash inhalation. Part I: filtration efficiency tests. (https://www.sciencedirect.com/science/article/pii/S1438463917308003?via%3Dihub) Int J Hyg Environ Health 2018 July;221(6):967-76
16. **Bowen LE.** Does that face mask really protect you? (https://journals.sagepub.com/doi/abs/10.1177/153567601001500204) Appl Biosaf 2010 Jun 1;15(2):67-71
17. **Shakya KM, Noyes A, Kallin R, et al.** Evaluating the efficacy of cloth facemasks in reducing particulate matter exposure. (https://www.nature.com/articles/jes201642) J Expo Sci Environ Epidemiol 2017 May;27(3):352-7
18. **van der Sande M., Teunis P, Sabel R.** Professional and home-made face masks reduce exposure to respiratory infections among the general population (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0002618) . PLOS One 2008 Jul 9;3(7):0002618
19. **Derrick JL, Gomersall CD.** Protecting healthcare staff from severe acute respiratory syndrome: filtration capacity of multiple surgical masks

(https://www.journalofhospitalinfection.com/article/S0195-6701(04)00479-7/abstract) . J Hosp Infect 2005 Apr;59(4):365-8

20. **Chughtai AA, Seale H, MacIntyre CR.** Use of cloth masks in the practice of infection control—evidence and policy gaps (https://www.ijic.info/article/view/11366) . Int J Infect Control 2013 Jun;9(3)
21. **Kellogg WH, MacMillan G.** An experimental study of the efficacy of gauze face masks. (https://ajph.aphapublications.org/doi/10.2105/AJPH.10.1.34) Am J Public Health 1920;10(1):34-42
22. **Saunders-Hastings P, Crispo JA, Sikora L, et al.** Effectiveness of personal protective measures in reducing pandemic influenza transmission: A systematic review and meta-analysis. (https://www.sciencedirect.com/science/article/pii/S1755436516300858?via%3Dihub) Epidemics 2017 Sep;20:1-20
23. **Cowling B J, Zhou Y, Ip DKM, et al.** Face masks to prevent transmission of influenza virus: a systematic review. (https://www.cambridge.org/core/journals/epidemiology-and-infection/article/face-masks-to-prevent-transmission-of-influenza-virus-a-systematic-review/64D368496EBDE0AFCC6639CCC9D8BC05) Epidemiol Infect 2010 Jan 22;138(4):449-56
24. **Canini L, Andreoletti L, Ferrari P, et al.** Surgical mask to prevent influenza transmission in households: a cluster randomized trial (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0013998) . PLOS One 2010 Nov 17;5(11):e13998
25. **MacIntyre CR, Zhang Y, Chughtai AA, et al.** Cluster randomised controlled trial to examine medical mask use as source control for people with respiratory illness. (https://bmjopen.bmj.com/content/6/12/e012330.long) BMJ Open 2016 Dec 30;6(12):e012330
26. **Meleny FL.** Infection in clean operative wounds: a nine year study. Surg Gynecol Obstet 1935;60:264-75
27. **Orr NWM.** Is a mask necessary in the operating theater? Ann R Coll Surg Engl 1981;63:390-2
28. **Mitchell NJ, Hunt S.** Surgical face masks in modern operating rooms—a costly and unnecessary ritual? (https://www.journalofhospitalinfection.com/article/0195-6701(91)90148-2/pdf) J Hosp Infect 1991;18(3):239-42
29. **Tunevall TG.** Postoperative wound infections and surgical face masks: a controlled study (https://link.springer.com/article/10.1007/BF01658736) . World J Surg 1991 May-Jun;15(3):383-7
30. **Belkin NL.** Masks, barriers, laundering, and gloving: Where is the evidence? (https://aornjournal.onlinelibrary.wiley.com/doi/abs/10.1016/S0001-2092%2806%2963946-X) AORN J 2006 Oct 25;84(4):655-63
31. **Johnson DF, Druce JD, Birch C, et al.** A quantitative assessment of the efficacy of surgical and N95 masks to filter influenza virus in patients with acute influenza infection. (https://academic.oup.com/cid/article/49/2/275/405108) Clin Infect Dis 2009 Jul 15;49(2):275-7
32. **Driessche KV, Hens N, Tilley P, et al.** Surgical masks reduce airborne spread of Pseudomonas aeruginosa in colonized patients with cystic fibrosis. (https://www.atsjournals.org/doi/full/10.1164/rccm.201503-0481LE?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub%3Dpubmed&) Am J Respir Crit Care Med 2015 Oct 1;192(7):897-9
33. **Milton DK, Fabian MP, Cowling BJ, et al.** Influenza virus aerosols in human exhaled breath: particle size, culturability, and effect of surgical masks. (https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1003205) PLoS Pathog 2013 Mar;9(3):e1003205
34. **Stockwell RE, Wood ME, He C, et al.** Face masks reduce the release of Pseudomonas aeruginosa cough aerosols when worn for clinically relevant periods.

(https://www.atsjournals.org/doi/full/10.1164/rccm.201805-0823LE?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3dpubmed) Am J Respir Crit Care Med 2018 Nov 15;198(10):1339-42

35. **Hui DS, Chow BK, Chu L, et al.** Exhaled air dispersion during coughing with and without wearing a surgical or N95 mask. (http://dx.plos.org/10.1371/journal.pone.0050845) PloS One 2012;7(12)e50845
36. **Dharmadhikari AS, Mphahlele M, Stoltz A, et al.** Surgical face masks worn by patients with multidrug-resistant tuberculosis: impact on infectivity of air on a hospital ward. (https://www.atsjournals.org/doi/full/10.1164/rccm.201107-1190OC?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3dpubmed) Am J Respir Crit Care Med 2012 May 15;185(10):1104-9
37. **Sung AD, Sung JA, Thomas S, et al.** Universal mask usage for reduction of respiratory viral infections after stem cell transplant: a prospective trial. (https://academic.oup.com/cid/article/63/8/999/2389110) Clin Infect Dis 2016 Oct 15;63(8):999-1006
38. **MacIntyre CR, Seale H, Dung TC, et al.** A cluster randomised trial of cloth masks compared with medical masks in healthcare workers. (https://bmjopen.bmj.com/content/5/4/e006577.long) BMJ Open 2015 Apr 22;5(4):e006577
39. **Loeb M, Dafoe N, Mahony J, et al.** Surgical mask vs N95 respirator for preventing influenza among healthcare workers: a randomized trial. (https://jamanetwork.com/journals/jama/article-abstract/184819) JAMA 2009 Nov 4;302(17):1865-71
40. **MacIntyre CR, Wang Q, Cauchemez S, et al.** A cluster randomized clinical trial comparing fit-tested and non-fit-tested N95 respirators to medical masks to prevent respiratory virus infection in health care workers (https://onlinelibrary.wiley.com/doi/full/10.1111/j.1750-2659.2011.00198.x) . Influenza Other Respir Viruses 2011;5(3):170-9
41. **MacIntyre CR, Wang Q, Rahman B, et al.** Efficacy of face masks and respirators in preventing upper respiratory tract bacterial colonization and co-infection in hospital healthcare workers—authors' reply (https://www.sciencedirect.com/science/article/pii/S009174351400190X) . Prev Med 2014 Aug;65:154
42. **MacIntyre CR, Wang Q, Seale H, et al.** A randomized clinical trial of three options for N95 respirators and medical masks in health workers (https://www.atsjournals.org/doi/full/10.1164/rccm.201207-1164OC) . Am J Resp Crit Care Med 2013;187(9):960-6
43. **Radonovich LJ, Simberkoff MS, Bessesen MT, et al.** N95 respirators vs medical masks for preventing influenza among health care personnel: a randomized clinical trial. (https://jamanetwork.com/journals/jama/article-abstract/2749214) JAMA 2019 Sep 3;322(9):824-33
44. **Gralton J, and McLaws ML.** Protecting healthcare workers from pandemic influenza: N95 or surgical masks?. (https://journals.lww.com/ccmjournal/Abstract/2010/02000/Protecting_healthcare_workers_from_pandemic.40.aspx) Crit Care Med 2010 Feb;38(2):657-67
45. **bin-Reza F, Chavarrias VL, Nicoll A, et al.** The use of masks and respirators to prevent transmission of influenza: a systematic review of the scientific evidence (https://onlinelibrary.wiley.com/doi/full/10.1111/j.1750-2659.2011.00307.x) . Influenza Other Respir Virus 2012 Jul;6(4):257-67
46. **Bunyan D, Ritchie L, Jenkins D, et al.** Respiratory and facial protection: a critical review of recent literature. (https://linkinghub.elsevier.com/retrieve/pii/S0195-6701(13)00280-6) J Hosp Infect 2013

Nov;85(3):165-9
47. **Smith JD, MacDougall CC, Johnstone J, et al.** Effectiveness of N95 respirators versus surgical masks in protecting health care workers from acute respiratory infection: a systematic review and meta-analysis. (https://www.cmaj.ca/content/188/8/567.long) CMAJ 2016 May 17;188(8):567-74
48. **Jefferson T, Jones M, Ansari LAA, et al.** Physical interventions to interrupt or reduce the spread of respiratory viruses. Part 1 - Face masks, eye protection and person distancing: systematic review and meta-analysis. (https://www.medrxiv.org/content/10.1101/2020.03.30.20047217v1) medRxiv 2020 Mar 30
49. **Offeddu V, Yung CF, Low MSF, et al.** Effectiveness of masks and respirators against respiratory infections in healthcare workers: a systematic review and meta-analysis. (https://academic.oup.com/cid/article/65/11/1934/4068747) Clin Infect Dis 2017 Aug 7;65(11):1934-42
50. **Long Y, Hu T, Liu L, et al.** Effectiveness of N95 respirators versus surgical masks against influenza: A systematic review and meta-analysis. (https://onlinelibrary.wiley.com/doi/full/10.1111/jebm.12381) J Evid Based Med 2020 (published online Mar 13)
51. **MacIntyre CR, Chughtai AA, Rahman B, et al.** The efficacy of medical masks and respirators against respiratory infection in healthcare workers (https://onlinelibrary.wiley.com/doi/full/10.1111/irv.12474) . Influenza Other Respir Viruses 2017;11(6):511-7
52. **Loeb M, McGeer A, Henry B, et al.** SARS among critical care nurses, Toronto. (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3322898/) Emerg Infect Dis 2004 Feb;10(2):251-5

RELATED NEWS

ALL NEWS

SEP 02 2020 **Health agency: COVID-19 hitting health workers hard in Americas**

SEP 02 2020 **Studies, WHO back steroid use in very ill COVID-19 patients**

SEP 02 2020 **Iowa emerges as national COVID-19 hot spot**

SEP 01 2020 **19-month pandemic-related surgical backlog in Ontario**

CIDRAP - Center for Infectious Disease Research and Policy
Office of the Vice President for Research, University of Minnesota, Minneapolis, MN

ScienceDaily®

Your source for the latest research news

One Virus Particle Is Enough To Cause Infectious Disease

Date: March 14, 2009

Source: Wageningen University and Research Centre

Summary: Can exposure to a single virus particle lead to infection or disease? Until now, solid proof has been lacking. Experimental research with insect larvae has shown that one virus particle is theoretically enough to cause infection and subsequent disease.

FULL STORY

Can exposure to a single virus particle lead to infection or disease? Until now, solid proof has been lacking. Experimental research with insect larvae at Wageningen University and Simon Fraser University in Canada has shown that one virus particle is theoretically enough to cause infection and subsequent disease.

A virus population is usually composed of a collection of variants of virus particles. In order to investigate whether virus particles (virions) can cause an infection independently from each other, and therefore individually, the researchers set up an experiment with two 'marked' virus variants. They exposed a population of hosts (caterpillars) to both variants.

The experiment showed that exposure to a low dosage of virus particles resulted in a small number host infections (20%). The majority of these hosts (86%) turned out to be infected by a single virus genotype. In contrast, exposure to a high dosage of virus particles resulted in virtually all the hosts (99%) becoming infected, where most of the hosts were infected by both types of virus. Only 14% were infected by only one of the two variants.

Based on the assumption that every virus particle operates independently from all other virus particles, the researchers set up a probability model. This model predicts how many virus particles have caused an infection and how many different virus genotypes are present in infected hosts, such as plants, insects or people. The results of the infection experiment with the susceptible insects are in agreement with the model predictions. From this it can be derived that the virus particles have an independent effect, and that a single virus particle can indeed cause infection and/or disease.

If there are few virus particles that lead to an infection, the number of virus particles determines the degree of diversity that can be present within the host. This is an important finding because the interactions between virus variants, such as competition and exchanging genetic information, determine the progression of disease and the evolution of the virus.

Until now, it was unclear whether a virus must be seen as an individual that can infect a host independently, or whether a cloud of viruses 'cooperates' to cause an infection. It is not yet known if the viruses that affect people can also act individually, but this research shows that it is possible.

The researchers recently published this finding in the *Proceedings of the Royal Society B.*

Story Source:

Materials provided by **Wageningen University and Research Centre**. *Note: Content may be edited for style and length.*

Journal Reference:

1. Mark P Zwart, Lia Hemerik, Jenny S Cory, J. Arjan G.M de Visser, Felix J.J.A Bianchi, Monique M Van Oers, Just M Vlak, Rolf F Hoekstra, and Wopke Van der Werf. **An experimental test of the independent action hypothesis in virus%u2013insect pathosystems**. *Proc. R. Soc. B*, 2009; DOI: 10.1098/rspb.2009.0064

Cite This Page:

MLA APA Chicago

Wageningen University and Research Centre. "One Virus Particle Is Enough To Cause Infectious Disease." ScienceDaily. ScienceDaily, 14 March 2009. <www.sciencedaily.com/releases/2009/03/090313150254.htm>.

RELATED STORIES

Seeing a Virus in Action

Aug. 14, 2017 — Imaging the movement of a virus demonstrates that single-particle X- ray scattering has the potential to shed new light on key molecular processes, like viral infection, when paired with powerful new ...

Epstein-Barr Virus and Cancer: New Tricks from an Old Dog

Feb. 13, 2017 — Approximately 98 percent of all adults are infected with the Epstein-Barr virus. In rare cases, an infection with this virus causes cancer. Scientists have now discovered that a component of the ...

Monkey Study Shows Zika Infection Prolonged in Pregnancy

June 28, 2016 — Researchers studying monkeys have shown that one infection with Zika virus protects against future infection, though pregnancy may drastically prolong the time the virus stays in the ...

Experimental Ebola Antibody Protects Monkeys

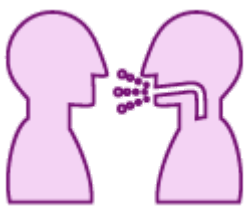
Feb. 25, 2016 — Scientists have discovered that a single monoclonal antibody isolated from a human Ebola virus disease survivor protected non-human primates when given as late as five days after lethal Ebola ...

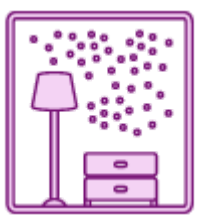
Centers for Disease Control and Prevention

Measles (Rubeola)

Transmission of Measles

Measles is highly contagious. It spreads when an infected person coughs or sneezes.





Measles is a highly contagious virus that lives in the nose and throat mucus of an infected person. It can spread to others through coughing and sneezing. Also, measles virus can **live for up to two hours in an airspace where the infected person coughed or sneezed.**

If other people **breathe the contaminated air or touch the infected surface**, then touch their eyes, noses, or mouths, they can become infected. Measles is so contagious that if one person has it, up to 90% of the people close to that person who are not immune will also become infected.

Infected people can **spread measles to others from four days before through four days after the rash appears**.

Measles is a disease of humans; measles virus is not spread by any other animal species.

Call your doctor immediately if you think you or your child have been exposed.

Page last reviewed: February 5, 2018

generated droplets and to qualitatively describe the effect of a damp cloth cover over the mouth to curb the emission of droplets.

Philip Anfinrud, Ph.D.
Valentyn Stadnytskyi, Ph.D.
National Institutes of Health
Bethesda, MD

Christina E. Bax, B.A.
Perelman School of Medicine at the University of Pennsylvania
Philadelphia, PA

Adriaan Bax, Ph.D.
National Institutes of Health
Bethesda, MD

Disclosure forms provided by the authors are available with the full text of this letter at NEJM.org.

This letter was published on April 15, 2020, at NEJM.org.

1. Duguid JP. The size and the duration of air-carriage of respiratory droplets and droplet-nuclei. J Hyg (Lond) 1946;44:471-9.
2. Marr LC, Tang JW, Van Mullekom J, Lakdawala SS. Mechanistic insights into the effect of humidity on airborne influenza virus survival, transmission and incidence. J R Soc Interface 2019;16(150).
3. Asadi S, Wexler AS, Cappa CD, Barreda S, Bouvier NM, Ristenpart WD. Aerosol emission and superemission during human speech increase with voice loudness. Sci Rep 2019;9:2348.
4. Chao CYH, Wan MP, Morawska L, et al. Characterization of expiration air jets and droplet size distributions immediately at the mouth opening. J Aerosol Sci 2009;40:122-33.

DOI: 10.1056/NEJMc2007800

Droplets and Aerosols in the Transmission of SARS-CoV-2

TO THE EDITOR: Anfinrud et al. now illustrate in the *Journal*[1] how liquid droplets exhaled during speech can linger in the air. The large particles to which they refer remain airborne only briefly before settling because of gravity; these particles may pose a threat of infection if they are inhaled by persons close by as well as a contact hazard if they are transferred to another person's nasal or oral passages. In this way, persons infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) may contribute to the spread of the infection.

Breathing and talking also produce smaller and much more numerous particles, known as aerosol particles, than those visualized in the laser experiment of Anfinrud and colleagues.[2-4] Certain persons called "super spreaders" produce many more aerosol particles than other persons. The diameters of these particles are in the micron range. These particles are too small to settle because of gravity, but they are carried by air currents and dispersed by diffusion and air turbulence.

Inhaled droplets and aerosol particles have different sites of deposition in the recipient. Inhaled droplets are deposited in the upper regions of the respiratory tract, from which they may be removed in nasal secretions or carried upward by the mucociliary escalator, to be expelled or swallowed. In contrast, inhaled aerosolized particles can penetrate to the depths of the lungs, where they may be deposited in the alveoli.

A recent study, the results of which were also published in the *Journal*, showed that experimentally produced aerosols containing SARS-CoV-2 virions remained infectious in tissue-culture assays, with only a slight reduction in infectivity during a 3-hour period of observation.[5] Aerosols from infected persons may therefore pose an inhalation threat even at considerable distances and in enclosed spaces, particularly if there is poor ventilation. The possible contribution of infective aerosols to the current pandemic suggests the advisability of wearing a suitable mask whenever it is thought that infected persons may be nearby and of providing adequate ventilation of enclosed spaces where such persons are known to be or may recently have been.

Matthew Meselson, Ph.D.
Harvard University
Cambridge, MA
msm@wjh.harvard.edu

Disclosure forms provided by the author are available with the full text of this letter at NEJM.org.

This letter was published on April 15, 2020, at NEJM.org.

1. Anfinrud P, Stadnytskyi V, Bax CE, Bax A. Visualizing speech-generated oral fluid droplets with laser light scattering. N Engl J Med 2020;382:2061-2.
2. Edwards DA, Man JC, Brand P, et al. Inhaling to mitigate exhaled bioaerosols. Proc Natl Acad Sci U S A 2004;101:17383-8.
3. Asadi S, Wexler AS, Cappa CD, Barreda S, Bouvier NM, Ristenpart WD. Aerosol emission and superemission during human speech increase with voice loudness. Sci Rep 2019;9:2348.
4. Tellier R, Li Y, Cowling BJ, Tang JW. Recognition of aerosol transmission of infectious agents: a commentary. BMC Infect Dis 2019;19:101.
5. van Doremalen N, Bushmaker T, Morris DH, et al. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. N Engl J Med 2020;382:1564-7.

DOI: 10.1056/NEJMc2009324

generated droplets and to qualitatively describe the effect of a damp cloth cover over the mouth to curb the emission of droplets.

Philip Anfinrud, Ph.D.
Valentyn Stadnytskyi, Ph.D.
National Institutes of Health
Bethesda, MD

Christina E. Bax, B.A.
Perelman School of Medicine at the University of Pennsylvania
Philadelphia, PA

Adriaan Bax, Ph.D.
National Institutes of Health
Bethesda, MD

Disclosure forms provided by the authors are available with the full text of this letter at NEJM.org.

This letter was published on April 15, 2020, at NEJM.org.

1. Duguid JP. The size and the duration of air-carriage of respiratory droplets and droplet-nuclei. J Hyg (Lond) 1946;44:471-9.
2. Marr LC, Tang JW, Van Mullekom J, Lakdawala SS. Mechanistic insights into the effect of humidity on airborne influenza virus survival, transmission and incidence. J R Soc Interface 2019;16(150).
3. Asadi S, Wexler AS, Cappa CD, Barreda S, Bouvier NM, Ristenpart WD. Aerosol emission and superemission during human speech increase with voice loudness. Sci Rep 2019;9:2348.
4. Chao CYH, Wan MP, Morawska L, et al. Characterization of expiration air jets and droplet size distributions immediately at the mouth opening. J Aerosol Sci 2009;40:122-33.

DOI: 10.1056/NEJMc2007800

Droplets and Aerosols in the Transmission of SARS-CoV-2

TO THE EDITOR: Anfinrud et al. now illustrate in the *Journal*[1] how liquid droplets exhaled during speech can linger in the air. The large particles to which they refer remain airborne only briefly before settling because of gravity; these particles may pose a threat of infection if they are inhaled by persons close by as well as a contact hazard if they are transferred to another person's nasal or oral passages. In this way, persons infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) may contribute to the spread of the infection.

Breathing and talking also produce smaller and much more numerous particles, known as aerosol particles, than those visualized in the laser experiment of Anfinrud and colleagues.[2-4] Certain persons called "super spreaders" produce many more aerosol particles than other persons. The diameters of these particles are in the micron range. These particles are too small to settle because of gravity, but they are carried by air currents and dispersed by diffusion and air turbulence.

Inhaled droplets and aerosol particles have different sites of deposition in the recipient. Inhaled droplets are deposited in the upper regions of the respiratory tract, from which they may be removed in nasal secretions or carried upward by the mucociliary escalator, to be expelled or swallowed. In contrast, inhaled aerosolized particles can penetrate to the depths of the lungs, where they may be deposited in the alveoli.

A recent study, the results of which were also published in the *Journal*, showed that experimentally produced aerosols containing SARS-CoV-2 virions remained infectious in tissue-culture assays, with only a slight reduction in infectivity during a 3-hour period of observation.[5] Aerosols from infected persons may therefore pose an inhalation threat even at considerable distances and in enclosed spaces, particularly if there is poor ventilation. The possible contribution of infective aerosols to the current pandemic suggests the advisability of wearing a suitable mask whenever it is thought that infected persons may be nearby and of providing adequate ventilation of enclosed spaces where such persons are known to be or may recently have been.

Matthew Meselson, Ph.D.
Harvard University
Cambridge, MA
msm@wjh.harvard.edu

Disclosure forms provided by the author are available with the full text of this letter at NEJM.org.

This letter was published on April 15, 2020, at NEJM.org.

1. Anfinrud P, Stadnytskyi V, Bax CE, Bax A. Visualizing speech-generated oral fluid droplets with laser light scattering. N Engl J Med 2020;382:2061-2.
2. Edwards DA, Man JC, Brand P, et al. Inhaling to mitigate exhaled bioaerosols. Proc Natl Acad Sci U S A 2004;101:17383-8.
3. Asadi S, Wexler AS, Cappa CD, Barreda S, Bouvier NM, Ristenpart WD. Aerosol emission and superemission during human speech increase with voice loudness. Sci Rep 2019;9:2348.
4. Tellier R, Li Y, Cowling BJ, Tang JW. Recognition of aerosol transmission of infectious agents: a commentary. BMC Infect Dis 2019;19:101.
5. van Doremalen N, Bushmaker T, Morris DH, et al. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. N Engl J Med 2020;382:1564-7.

DOI: 10.1056/NEJMc2009324

Title:

Closed environments facilitate secondary transmission of coronavirus disease 2019 (COVID-19)

Running title: Closed environment and COVID-19

Authors:

Hiroshi Nishiura, M.D., Ph.D.[1, 2], Hitoshi Oshitani, M.D., Ph.D. [1, 3], Tetsuro Kobayashi, M.D. [1, 2], Tomoya Saito, M.D., M.P.H., Ph.D. [1, 4], Tomimasa Sunagawa, M.D., Ph.D. [1, 5], Tamano Matsui, M.D., Ph.D. [1, 5], Takaji Wakita, M.D., Ph.D. [1, 5], MHLW COVID-19 Response Team[1], Motoi Suzuki, M.D., Ph.D. [1, 5]

1. Cluster Intervention Group, Ministry of Health, Labour and Welfare, Kasumigaseki 1-2-2, Chiyoda-ku, Tokyo 100-8916, Japan

2. Graduate School of Medicine, Hokkaido University, Kia 15 Jo Nishi 7 Chome, Kitaku, Sapporo, 060-8638, Japan

3. Tohoku University Graduate School of Medicine, Seiryocho 2-1, Aoba-ku, Sendai, 980-8575, Japan

4. National Institute of Public Health, 5. National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan

(Correspondence to Hiroshi Nishiura at: Address: Kita 15 Jo Nishi 7 Chome, Kita-ku, Sapporo-shi, Hokkaido 060-8638, Japan, Tel: +81-11-706-5066; Fax: +81-11-706-7819; Email: nishiurah@med.hokudai.ac.jp)

Abstract

Objective: To identify common features of cases with novel coronavirus disease (COVID-19) so as to better understand what factors promote secondary transmission including superspreading events.

Methods: A total of 110 cases were examined among eleven clusters and sporadic cases, and investigated who acquired infection from whom. The clusters included four in Tokyo and one each in Aichi, Fukuoka, Hokkaido, Ishikawa, Kanagawa and Wakayama prefectures. The number of secondary cases generated by each primary case was calculated using contact tracing data.

Results: Of the 110 cases examined, 27 (24.6%) were primary cases who generated secondary cases. The odds that a primary case transmitted COVID-19 in a closed environment was 18.7 times greater compared to an open-air environment (95% confidence interval [CI]: 6.0, 57.9).

Conclusions: It is plausible that closed environments contribute to secondary transmission of COVID-19 and promote superspreading events. Our findings are also consistent with the declining incidence of COVID-19 cases in China, as gathering in closed environments was prohibited in the wake of the rapid spread of the disease.

Introduction

Although the incidence of coronavirus disease 2019 (COVID-19) in China began to decrease in February 2020,[1] many countries are struggling with containment of the disease. To effectively reduce the spread of COVID-19, it is vital to identify common features of cases so as to better understand what factors promote superspreading events,[2] wherein an extraordinarily large number of secondary transmissions are produced by a single primary case. Commissioned by the Minister of the Ministry of Health, Labour, and Welfare of Japan, we collected secondary transmission data with the aim of identifying high risk transmission settings.

Methods

As of 28 February 2020,[3] we examined a total of 110 cases among eleven clusters and sporadic cases, and investigated who acquired infection from whom. The clusters included four in Tokyo and one each in Aichi, Fukuoka, Hokkaido, Ishikawa, Kanagawa and Wakayama prefectures. All traced transmission events were examined in relation to close contact in indoor environments, including fitness gyms, a restaurant boat on a river, hospitals, and a snow festival where there were eating spaces in tents with minimal ventilation rate. The number of secondary cases generated by each primary case was calculated using contact tracing data.

Results

Of the 110 cases examined, 27 (24.6%) were primary cases who generated secondary cases. Figure 1 shows the distribution of these transmissions, of which the mean and variance were 0.6 cases and 2.5 cases^2, respectively. The odds that a primary

case transmitted COVID-19 in a closed environment was 18.7 times greater compared to an open-air environment (95% confidence interval [CI]: 6.0, 57.9).

If superspreading events are defined as events where the number of secondary cases generated by a single primary case is greater than the 95th percentile of the distribution (i.e. transmission to three or more persons), then seven of the 110 cases (6.4%) were involved in such events. Six of these events (85.7%) took place in closed environments, and the odds ratio (OR) of superspreading events in closed environments was as high as 32.6 (95% CI: 3.7, 289.5).

Discussion

It is plausible that closed environments contribute to secondary transmission of COVID-19 and promote superspreading events. Closed environments are consistent with environmental sampling study[4] and also large-scale COVID-19 transmission events such as that of the ski chalet-associated cluster in France and the church- and hospital-associated clusters in South Korea[5]. Our findings are also consistent with the declining incidence of COVID-19 cases in China, as gathering in closed environments was prohibited in the wake of the rapid spread of the disease.

Reduction of unnecessary close contact in closed environments may help prevent large case clusters and superspreading events. We hope that with such a reduction in contact the reproduction number of COVID-19 in Japan will be maintained below 1 and contact tracing will be sufficient to contain disease spread.[6] As the possibility of confounders and interactions was not assessed in this study, additional studies must be conducted to verify the importance of closed environments as facilitators for transmission of COVID-19.

Conflict of interest:

We declare that we have no conflict of interest.

Acknowledgement:

We sincerely thank staff of local governments, including health centers and prefectural institutes of public health, healthcare facilities, and associated companies and organizations for cooperating us to collect and investigate secondary transmission data. H.N. received funding support from Japan Agency for Medical Research and Development [grant number: JP18fk0108050] and the Japan Science and Technology Agency (JST) Core Research for Evolutional Science and Technology (CREST) program [grant number: JPMJCR1413].

References:

1. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. Vital Surveillances: The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) — China, 2020. China CDC Weekly 2020;2(8):113-122.

2. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. Nature 2005;438:355-359.

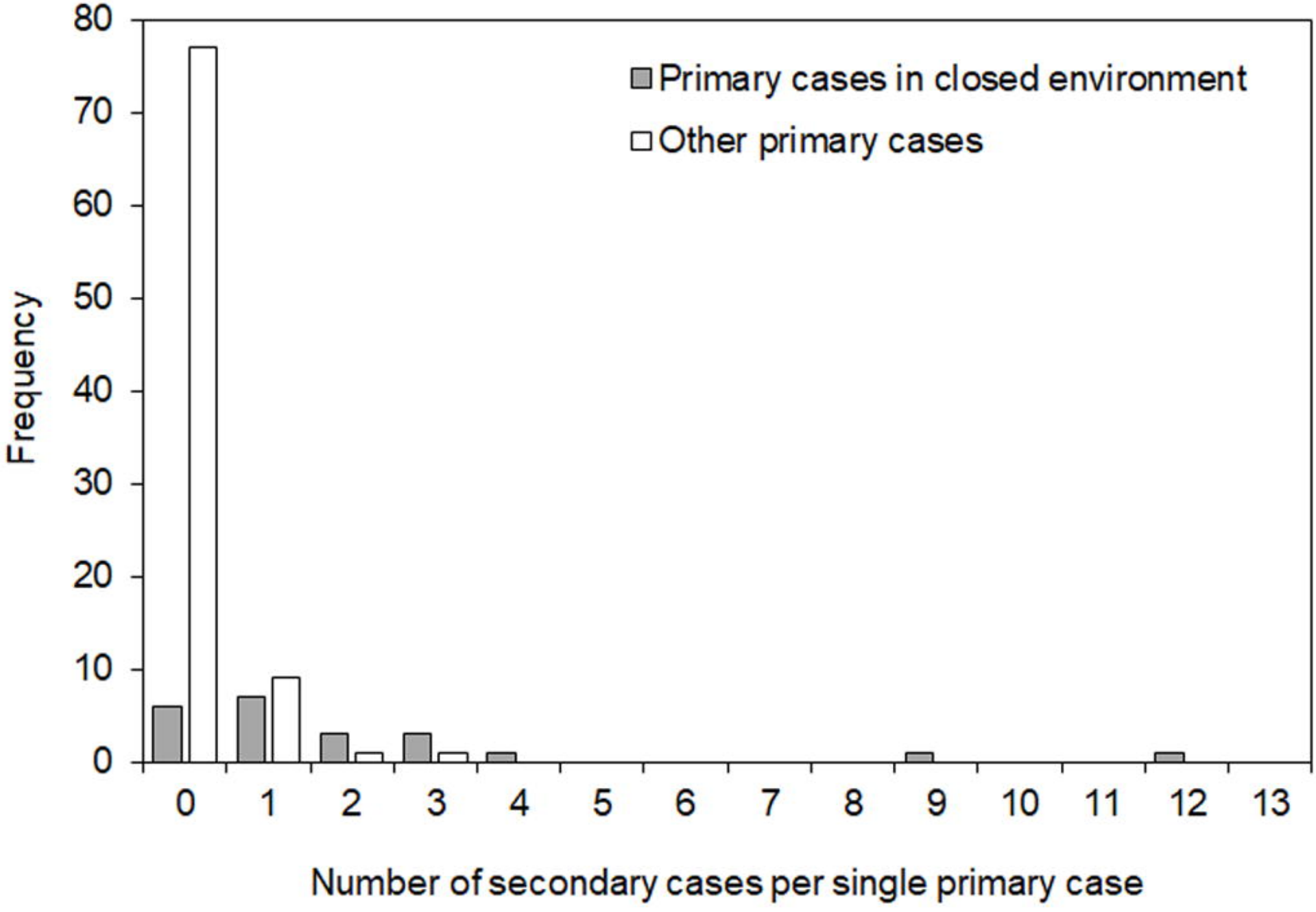
3. Ministry of Health, Labour and Welfare, Japan. On the novel coronavirus infection. Tokyo: Ministry of Health, Labour and Welfare, 2020. Available online from: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000164708_00001.html (accessed on 26 February 2020).

4. Ong SWX, Tan YK, Chia PY, Lee TH, Ng OT, Wong MSY, Marimuthu K. Air,

109 Surface Environmental, and Personal Protective Equipment Contamination by Severe
110 Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) From a Symptomatic
111 Patient. JAMA. 2020; in press. doi: 10.1001/jama.2020.3227.
112 5. Promed mail. Subject: PRO/AH/EDR> COVID-19 update (19): China, global, Italy
113 & Iran. imported cases, WHO. Archive Number: 20200226.7029842
114 6. Grantz K, Metcalf JE, Lessler J. Dispersion vs. control. 2020. Available online from:
115 https://hopkinsidd.github.io/nCoV-Sandbox/DispersionExploration.html (accessed on
116 26 February 2020).
117

118 ## Figure legend

119 Figure 1. The distribution of the number of secondary cases generated by a single
120 primary case with novel coronavirus (COVID-19). The mean and variance were 0.6
121 cases and 2.5 cases^2, respectively.


Primary cases in closed environment
Other primary cases
Frequency
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
Number of secondary cases per single primary case

Closed environments facilitate secondary transmission of coronavirus disease 2019 (COVID-19)

Hiroshi Nishiura, Hitoshi Oshitani, Tetsuro Kobayashi, Tomoya Saito, Tomimasa Sunagawa, Tamano Matsui, Takaji Wakita, MHLW COVID-19 Response Team, Motoi Suzuki

doi: https://doi.org/10.1101/2020.02.28.20029272

Abstract

Objective: To identify common features of cases with novel coronavirus disease (COVID-19) so as to better understand what factors promote secondary transmission including superspreading events. Methods: A total of 110 cases were examined among eleven clusters and sporadic cases, and investigated who acquired infection from whom. The clusters included four in Tokyo and one each in Aichi, Fukuoka, Hokkaido, Ishikawa, Kanagawa and Wakayama prefectures. The number of secondary cases generated by each primary case was calculated using contact tracing data. Results: Of the 110 cases examined, 27 (24.6%) were primary cases who generated secondary cases. The odds that a primary case transmitted COVID-19 in a closed environment was 18.7 times greater compared to an open-air environment (95% confidence interval [CI]: 6.0, 57.9). Conclusions: It is plausible that closed environments contribute to secondary transmission of COVID-19 and promote superspreading events. Our findings are also consistent with the declining incidence of COVID-19 cases in China, as gathering in closed environments was prohibited in the wake of the rapid spread of the disease.

Competing Interest Statement

The authors have declared no competing interest.

Funding Statement

H.N. received funding from the Japan Agency for Medical Research and Development (AMED) [grant number: JP18fk0108050]; the Japan Society for the Promotion of Science (JSPS) KAKENHI [grant numbers, H.N.: 17H04701, 17H05808, 18H04895 and 19H01074], the Inamori Foundation, and the Japan Science and Technology Agency (JST) CREST program [grant number: JPMJCR1413].

Author Declarations

All relevant ethical guidelines have been followed; any necessary IRB and/or ethics committee approvals have been obtained and details of the IRB/oversight body are included in the manuscript.

Yes

All necessary patient/participant consent has been obtained and the appropriate institutional forms have been archived.

Yes

I understand that all clinical trials and any other prospective interventional studies must be registered with an ICMJE-approved registry, such as ClinicalTrials.gov. I confirm that any such study reported in the manuscript has been registered and the trial registration ID is provided (note: if posting a prospective study registered retrospectively, please provide a statement in the trial ID field explaining why the study was not registered in advance).

Yes

I have followed all appropriate research reporting guidelines and uploaded the relevant EQUATOR Network research reporting checklist(s) and other pertinent material as supplementary files, if applicable.

Yes

Paper in collection COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv

---

Blog posts linking to this article:

Leiter Reports: A Philosophy Blog, 26 Aug 2020

Here. I guess I'm persuaded it's clearly a factor (much more than fomites), but it's not that easy, as the author explains via…

Booster Shots - latimes.com, 21 Aug 2020

One reason L.A.'s homeless people have avoided a COVID disaster could be that they live outside

R-bloggers, 22 Jul 2020

[This article was first published on Fabian Dablander, and kindly contributed to R-bloggers]. (You can report issue about…

PsyBlog, 18 Jul 2020

Can social distancing work or there are better ways to lower COVID-19 spread? → Support PsyBlog for just $5 per month. Enables…

Croakey, 06 Jul 2020

In her latest edition of The Health Wrap, Dr Lesley Russell investigates the latest pandemic news with a focus on health…

Iowa Climate, 05 Jul 2020

Another one reposted from the Cliff Mass Weather…

Climate Science, 05 Jul 2020

The Safety of Outdoor Air for #Coronavirus Is Now…

Whatts Up With That, 05 Jul 2020

There is now powerful observational evidence that outdoor air is extraordinary safe regarding COVID-19, and the recent protests…

Science news, comment and analysis | theguardian.com, 05 Jul 2020

The paediatrician and member of Independent Sage on Matt Hancock, the likelihood of a vaccine and why 50,000 deaths were…

Well - New York Times, 03 Jul 2020

With the virus raging in many parts of the country, new restrictions have left many wondering about the safety of a backyard…

N+1: научные статьи, новости, открытия, 02 Jul 2020

Наступило ли время второй волны COVID-19

Latest BMJ blogs, 30 Jun 2020

Deaths and new infections from covid-19 are falling to low levels in Scotland. Devi Sridhar and Adriel Chen look at how the…

Qubit, 29 Jun 2020

Több mint száz napja, hogy a SARS-CoV-2 koronavírus itthon is alaposan felforgatta a hétköznapjainkat. Sokan sokféleképpen éltük…

Indian Strategic Studies, 02 Jun 2020

Source LinkTO REALLY UNDERSTAND how the disease Covid-19 spreads, you have to see the world the way a virus moves through it. It…

Feed: All Latest, 28 May 2020

As public spaces reopen, scientists are racing to understand the mysterious and turbulent way the disease spreads through air—fro…

Back to top

Previous Next

Posted April 16, 2020.

COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv

Subject Area

Epidemiology

Subject Areas

All Articles

- Addiction Medicine
- Allergy and Immunology
- Anesthesia
- Cardiovascular Medicine
- Dentistry and Oral Medicine
- Dermatology
- Emergency Medicine
- Endocrinology (including Diabetes Mellitus and Metabolic Disease)
- Epidemiology
- Forensic Medicine
- Gastroenterology
- Genetic and Genomic Medicine
- Geriatric Medicine
- Health Economics
- Health Informatics
- Health Policy
- Health Systems and Quality Improvement
- Hematology
- HIV/AIDS

Infectious Diseases (except HIV/AIDS)

Intensive Care and Critical Care Medicine

Medical Education

Medical Ethics

Nephrology

Neurology

Nursing

Nutrition

Obstetrics and Gynecology

Occupational and Environmental Health

Oncology

Ophthalmology

Orthopedics

Otolaryngology

Pain Medicine

Palliative Medicine

Pathology

Pediatrics

Pharmacology and Therapeutics

Primary Care Research

Psychiatry and Clinical Psychology

Public and Global Health

Radiology and Imaging

Rehabilitation Medicine and Physical Therapy

Respiratory Medicine

Rheumatology

Sexual and Reproductive Health

Sports Medicine

Surgery

Toxicology

Transplantation

Urology

One Per Cent, 27 May 2020

Many countries are relaxing coronavirus restrictions. If you've been asked to return to work or school, how can you reduce the…

The Incidental Economist, 26 May 2020

The following originally appeared on the Baker Institute Blog and is coauthored by Vivian Ho, Ph.D. (@healthecontx), James A.

Booster Shots - latimes.com, 07 May 2020

The fight for access to open space and natural landscapes has a long history, and it's taken a new turn in the coronavirus…

Iowa Climate, 07 May 2020

by Judith…

Climate Etc., 07 May 2020

by Judith…

Science news, comment and analysis | theguardian.com, 04 May 2020

Dirty air is well known to worsen the heart and lung risk factors for Covid-19 - early research is cause for concernIn many…

DJG Blogger, 24 Apr 2020

From your lungs into the air around you, aerosols carry coronavirus. Peter Dazeley/The Image Bank via Getty…

Tweets referencing this article:

越後屋 もっち
@1905Tor

RT @Derive_ip: 忽那賢志は３密理論の根拠として下記論文の紹介の際にも「「3密空間」にいる感染者は、いない感染者よりも18.7倍も他の人へ感染させやすい」としていますが、そうは書いてないですよね。https://t.co/8an9k8eNHv 彼は紹介する論文をよく…

05:04AM

Samuel Stone
@SamThePol

@baby_no13 @politico Also read this, https://t.co/HOFQhAFtKR the pop density issue isn't the driver we all thought it was, probably because we misunderstood the spread: https://t.co/In8LQkMdya

07:20PM

medRxiv Comment Policy

Comments are moderated for offensive or irrelevant content (can take ~24 hours). Duplicated submission is unnecessary.

Please read our Comment Policy before commenting.

Everett • 4 months ago

Hey featheryHen,
here is the link you are asking for:
https://www.medrxiv.org/con...

Reply • Share ›

FeatheryHen • 4 months ago • edited

Interesting analysis, but I'm left wanting to know more about the source data. It would be good to se summary of the different transmission events you analysed and what the characteristics were, other than enclosed or not. Is a link to source data available?

2 Reply • Share ›

David Gordon Miller ➚ FeatheryHen • 4 months ago

https://www.nationalreview....

1 Reply • Share ›

FeatheryHen ➚ David Gordon Miller • 4 months ago

Thanks for replying but I'm really after the source data that the researchers are referri to. Often it's available as an appendix with final publication. I've been collecting info o case studies of indoor v outdoor transmission since this first hit the UK. Case studies where outdoor transmission has occurred seem to be very rare indeed.

3 Reply • Share ›

Donna Barstow Draws ➚ FeatheryHen • 3 months ago

Let me know if you find a source! Because I'm sure there is none. Outside indeed...

1 Reply • Share ›

BMJ Open

A cluster randomised trial of cloth masks compared with medical masks in healthcare workers

C Raina MacIntyre,[1] Holly Seale,[1] Tham Chi Dung,[2] Nguyen Tran Hien,[2] Phan Thi Nga,[2] Abrar Ahmad Chughtai,[1] Bayzidur Rahman,[1] Dominic E Dwyer,[3] Quanyi Wang[4]

To cite: MacIntyre CR, Seale H, Dung TC, *et al*. A cluster randomised trial of cloth masks compared with medical masks in healthcare workers. *BMJ Open* 2015;**5**: e006577. doi:10.1136/bmjopen-2014-006577

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2014-006577).

Received 9 September 2014
Revised 25 March 2015
Accepted 26 March 2015

For numbered affiliations see end of article.

Correspondence to
Professor C Raina MacIntyre; r.macintyre@unsw.edu.au

ABSTRACT

Objective: The aim of this study was to compare the efficacy of cloth masks to medical masks in hospital healthcare workers (HCWs). The null hypothesis is that there is no difference between medical masks and cloth masks.

Setting: 14 secondary-level/tertiary-level hospitals in Hanoi, Vietnam.

Participants: 1607 hospital HCWs aged ≥18 years working full-time in selected high-risk wards.

Intervention: Hospital wards were randomised to: medical masks, cloth masks or a control group (usual practice, which included mask wearing). Participants used the mask on every shift for 4 consecutive weeks.

Main outcome measure: Clinical respiratory illness (CRI), influenza-like illness (ILI) and laboratory-confirmed respiratory virus infection.

Results: The rates of all infection outcomes were highest in the cloth mask arm, with the rate of ILI statistically significantly higher in the cloth mask arm (relative risk (RR)=13.00, 95% CI 1.69 to 100.07) compared with the medical mask arm. Cloth masks also had significantly higher rates of ILI compared with the control arm. An analysis by mask use showed ILI (RR=6.64, 95% CI 1.45 to 28.65) and laboratory-confirmed virus (RR=1.72, 95% CI 1.01 to 2.94) were significantly higher in the cloth masks group compared with the medical masks group. Penetration of cloth masks by particles was almost 97% and medical masks 44%.

Conclusions: This study is the first RCT of cloth masks, and the results caution against the use of cloth masks. This is an important finding to inform occupational health and safety. Moisture retention, reuse of cloth masks and poor filtration may result in increased risk of infection. Further research is needed to inform the widespread use of cloth masks globally. However, as a precautionary measure, cloth masks should not be recommended for HCWs, particularly in high-risk situations, and guidelines need to be updated.

Trial registration number: Australian New Zealand Clinical Trials Registry: ACTRN12610000887077.

Strengths and limitations of this study

- The use of cloth masks is widespread around the world, particularly in countries at high-risk for emerging infections, but there have been no efficacy studies to underpin their use.
- This study is large, a prospective randomised clinical trial (RCT) and the first RCT ever conducted of cloth masks.
- The use of cloth masks are not addressed in most guidelines for health care workers—this study provides data to update guidelines.
- The control arm was 'standard practice', which comprised mask use in a high proportion of participants. As such (without a no-mask control), the finding of a much higher rate of infection in the cloth mask arm could be interpreted as harm caused by cloth masks, efficacy of medical masks, or most likely a combination of both.

INTRODUCTION

The use of facemasks and respirators for the protection of healthcare workers (HCWs) has received renewed interest following the 2009 influenza pandemic,[1] and emerging infectious diseases such as avian influenza,[2] Middle East respiratory syndrome coronavirus (MERS-coronavirus)[3 4] and Ebola virus.[5] Historically, various types of cloth/cotton masks (referred to here after as 'cloth masks') have been used to protect HCWs.[6] Disposable medical/surgical masks (referred to here after as 'medical masks') were introduced into healthcare in the mid 19th century, followed later by respirators.[7] Compared with other parts of the world, the use of face masks is more prevalent in Asian countries, such as China and Vietnam.[8–11]

In high resource settings, disposable medical masks and respirators have long since replaced the use of cloth masks in hospitals. Yet cloth masks remain widely used

globally, including in Asian countries, which have historically been affected by emerging infectious diseases, as well as in West Africa, in the context of shortages of personal protective equipment (PPE).[12 13] It has been shown that medical research disproportionately favours diseases of wealthy countries, and there is a lack of research on the health needs of poorer countries.[14] Further, there is a lack of high-quality studies around the use of facemasks and respirators in the healthcare setting, with only four randomised clinical trials (RCTs) to date.[15] Despite widespread use, cloth masks are rarely mentioned in policy documents,[16] and have never been tested for efficacy in a RCT. Very few studies have been conducted around the clinical effectiveness of cloth masks, and most available studies are observational or in vitro.[6] Emerging infectious diseases are not constrained within geographical borders, so it is important for global disease control that use of cloth masks be underpinned by evidence. The aim of this study was to determine the efficacy of cloth masks compared with medical masks in HCWs working in high-risk hospital wards, against the prevention of respiratory infections.

METHODS

A cluster-randomised trial of medical and cloth mask use for HCWs was conducted in 14 hospitals in Hanoi, Vietnam. The trial started on the 3 March 2011, with rolling recruitment undertaken between 3 March 2011 and 10 March 2011. Participants were followed during the same calendar time for 4 weeks of facemasks use and then one additional week for appearance of symptoms. An invitation letter was sent to 32 hospitals in Hanoi, of which 16 agreed to participate. One hospital did not meet the eligibility criteria; therefore, 74 wards in 15 hospitals were randomised. Following the randomisation process, one hospital withdrew from the study because of a nosocomial outbreak of rubella.

Participants provided written informed consent prior to initiation of the trial.

Randomisation

Seventy-four wards (emergency, infectious/respiratory disease, intensive care and paediatrics) were selected as high-risk settings for occupational exposure to respiratory infections. Cluster randomisation was used because the outcome of interest was respiratory infectious diseases, where prevention of one infection in an individual can prevent a chain of subsequent transmission in closed settings.[8 9] Epi info V.6 was used to generate a randomisation allocation and 74 wards were randomly allocated to the interventions.

From the eligible wards 1868 HCWs were approached to participate. After providing informed consent, 1607 participants were randomised by ward to three arms: (1) medical masks at all times on their work shift; (2) cloth masks at all times on shift or (3) control arm (standard practice, which may or may not include mask use). Standard practice was used as control because the IRB deemed it unethical to ask participants to not wear a mask. We studied continuous mask use (defined as wearing masks all the time during a work shift, except while in the toilet or during tea or lunch breaks) because this reflects current practice in high-risk settings in Asia.[8]

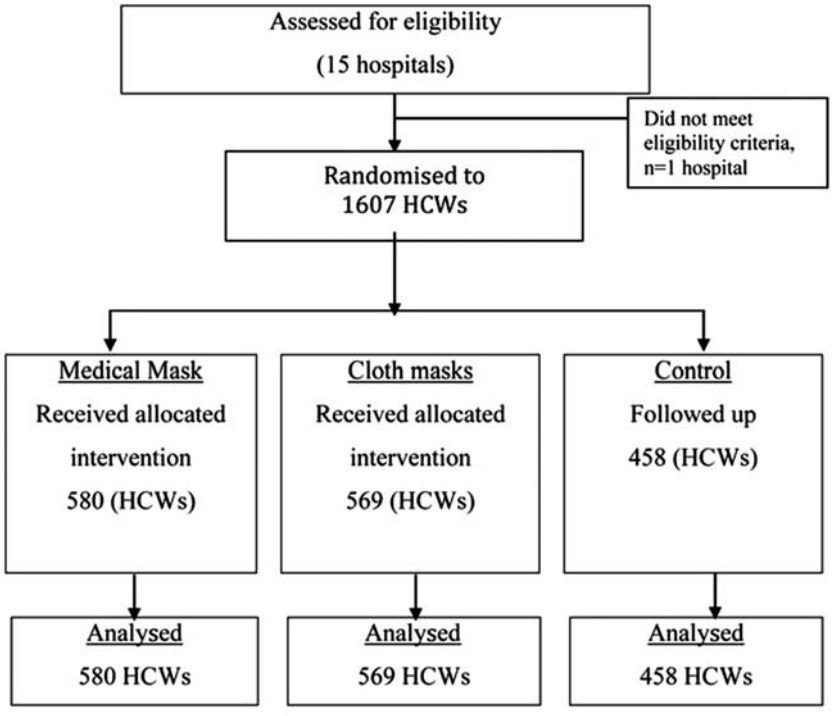


Figure 1 Consort diagram of recruitment and follow-up (HCWs, healthcare workers).

The laboratory results were blinded and laboratory testing was conducted in a blinded fashion. As facemask use is a visible intervention, clinical end points could not be blinded. Figure 1 outlines the recruitment and randomisation process.

Primary end points

There were three primary end points for this study, used in our previous mask RCTs:[8 9] (1) Clinical respiratory illness (CRI), defined as two or more respiratory symptoms or one respiratory symptom and a systemic symptom;[17] (2) influenza-like illness (ILI), defined as fever ≥38°C plus one respiratory symptom and (3) laboratory-confirmed viral respiratory infection. Laboratory confirmation was by nucleic acid detection using multiplex reverse transcriptase PCR (RT-PCR) for 17 respiratory viruses: respiratory syncytial virus (RSV) A and B, human metapneumovirus (hMPV), influenza A (H3N2), (H1N1)pdm09, influenza B, parainfluenza viruses 1–4, influenza C, rhinoviruses, severe acute respiratory syndrome (SARS) associated coronavirus (SARS-CoV), coronaviruses 229E, NL63, OC43 and HKU1, adenoviruses and human bocavirus (hBoV).[18–23] Additional end points included compliance with mask use, defined as using the mask during the shift for 70% or more of work shift hours.[9] HCWs were categorised as 'compliant' if the average use was equal or more than 70% of the working time. HCW were categorised as 'non-compliant' if the average mask use was less than 70% of the working time.

Eligibility

Nurses or doctors aged ≥18 years working full-time were eligible. Exclusion criteria were: (1) Unable or refused to consent; (2) Beards, long moustaches or long facial hair stubble; (3) Current respiratory illness, rhinitis and/or allergy.

Intervention

Participants wore the mask on every shift for four consecutive weeks. Participants in the medical mask arm were supplied with two masks daily for each 8 h shift, while participants in the cloth mask arm were provided with five masks in total for the study duration, which they were asked to wash and rotate over the study period. They were asked to wash cloth masks with soap and water every day after finishing the shifts. Participants were supplied with written instructions on how to clean their cloth masks. Masks used in the study were locally manufactured medical (three layer, made of non-woven material) or cloth masks (two layer, made of cotton) commonly used in Vietnamese hospitals. The control group was asked to continue with their normal practices, which may or may not have included mask wearing. Mask wearing was measured and documented for all participants, including the control arm.

Data collection and follow-up

Data on sociodemographic, clinical and other potential confounding factors were collected at baseline. Participants were followed up daily for 4 weeks (active intervention period), and for an extra week of standard practice, in order to document incident infection after incubation. Participants received a thermometer (traditional glass and mercury) to measure their temperature daily and at symptom onset. Daily diary cards were provided to record number of hours worked and mask use, estimated number of patient contacts (with/without ILI) and number/type of aerosol-generating procedures (AGPs) conducted, such as suctioning of airways, sputum induction, endotracheal intubation and bronchoscopy. Participants in the cloth mask and control group (if they used cloth masks) were also asked to document the process used to clean their mask after use.

We also monitored compliance with mask use by a previously validated self-reporting mechanism.[8] Participants were contacted daily to identify incident cases of respiratory infection. If participants were symptomatic, swabs of both tonsils and the posterior pharyngeal wall were collected on the day of reporting.

Sample collection and laboratory testing

Trained collectors used double rayon-tipped, plastic-shafted swabs to scratch tonsillar areas as well as the posterior pharyngeal wall of symptomatic participants. Testing was conducted using RT-PCR applying published methods.[19–23] Viral RNA was extracted from each respiratory specimen using the Viral RNA Mini kit (Qiagen, Germany), following the manufacturer's instructions. The RNA extraction step was controlled by amplification of a RNA house-keeping gene (amplify pGEM) using real-time RT-PCR. Only extracted samples with the house keeping gene detected by real-time RT-PCR were submitted for multiplex RT-PCR for viruses.

The reverse transcription and PCRs were performed in OneStep (Qiagen, Germany) to amplify viral target genes, and then in five multiplex RT-PCR: RSVA/B, influenza A/H3N2, A(H1N1) and B viruses, hMPV (reaction mix 1); parainfluenza viruses 1–4 (reaction mix 2); rhinoviruses, influenza C virus, SARS-CoV (reaction mix 3); coronaviruses OC43, 229E, NL63 and HKU1 (reaction mix 4); and adenoviruses and hBoV (reaction mix 5), using a method published by others.[18] All samples with viruses detected by multiplex RT-PCR were confirmed by virus-specific mono nested or hemi-nested PCR. Positive controls were prepared by in vitro transcription to control amplification efficacy and monitor for false negatives, and included in all runs (except for NL63 and HKU1). Each run always included two negatives to monitor amplification quality. Specimen processing, RNA extraction, PCR amplification and PCR product analyses were conducted in different rooms to avoid cross-contamination.[19 20]

Filtration testing

The filtration performance of the cloth and medical masks was tested according to the respiratory standard AS/NZS1716.[24] The equipment used was a TSI 8110 Filter tester. To test the filtration performance, the filter is challenged by a known concentration of sodium chloride particles of a specified size range and at a defined flow rate. The particle concentration is measured before and after adding the filter material and the relative filtration efficiency is calculated. We examined the performance of cloth masks compared with the performance levels—P1, P2 (=N95) and P3, as used for assessment of all particulate filters for respiratory protection. The 3M 9320 N95 and 3M Vflex 9105 N95 were used to compare against the cloth and medical masks.

Sample size calculation

To obtain 80% power at two-sided 5% significance level for detecting a significant difference of attack rate between medical masks and cloth masks, and for a rate of infection of 13% for cloth mask wearers compared with 6% in medical mask wearers, we would need eight clusters per arm and 530 participants in each arm, and intracluster correlation coefficient (ICC) 0.027, obtained from our previous study.[8] The design effect (deff) for this cluster randomisation trial was 1.65 (deff=1+(m −1)×ICC=1+(25−1)×0.027=1.65). As such, we aimed to recruit a sample size of 1600 participants from up to 15 hospitals.

Analysis

Descriptive statistics were compared among intervention and control arms. Primary end points were analysed by intention to treat. We compared the event rates for the primary outcomes across study arms and calculated p values from cluster-adjusted χ^2 tests[25] and ICC.[25 26] We also estimated relative risk (RR) after adjusting for clustering using a log-binomial model under generalised estimating equation (GEE) framework.[27] We checked for variables which were unequally distributed across arms, and conducted an adjusted analysis accordingly. We fitted a multivariable log-binomial model, using GEE to account for clustering by ward, to estimate RR after adjusting for potential confounders. In the initial model, we included all the variables that had p value less than 0.25 in the univariable analysis, along with the main exposure variable (randomisation arm). A backward elimination method was used to remove the variables that did not have any confounding effect.

As most participants in the control arm used a mask during the trial period, we carried out a post-hoc analysis comparing all participants who used only a medical mask (from the control arm and the medical mask arm) with all participants who used only a cloth mask (from the control arm and the cloth arm). For this analysis, controls who used both types of mask (n=245) or used N95 respirators (n=3) or did not use any masks (n=2) were excluded. We fitted a multivariable log-binomial model, to estimate RR after adjusting for potential confounders. As we pooled data of participants from all three arms and analysed by mask type, not trial arm, we did not adjust for clustering here. All statistical analyses were conducted using STATA V.12.[28]

Owing to a very high level of mask use in the control arm, we were unable to determine whether the differences between the medical and cloth mask arms were due to a protective effect of medical masks or a detrimental effect of cloth masks. To assist in interpreting the data, we compared rates of infection in the medical mask arm with rates observed in medical mask arms from two previous RCTs,[8 9] in which no efficacy of medical masks could be demonstrated when compared with control or N95 respirators, recognising that seasonal and geographic variation in virus activity affects the rates of exposure (and hence rates of infection outcomes) among HCWs. This analysis was possible because the trial designs were similar and the same outcomes were measured in all three trials. The analysis was carried out to determine if the observed results were explained by a detrimental effect of cloth masks or a protective effect of medical masks.

RESULTS

A total of 1607 HCWs were recruited into the study. The participation rate was 86% (1607/1868). The average number of participants per ward was 23 and the mean age was 36 years. On average, HCWs were in contact with 36 patients per day during the trial period (range 0–661 patients per day, median 20 patients per day). The distribution of demographic variables was generally similar between arms (table 1). Figure 2 shows the primary outcomes for each of the trial arms. The rates of CRI, ILI and laboratory-confirmed virus infections were lowest in the medical mask arm, followed by the control arm, and highest in the cloth mask arm.

Table 2 shows the intention-to-treat analysis. The rate of CRI was highest in the cloth mask arm, followed by the control arm, and lowest in the medical mask arm. The same trend was seen for ILI and laboratory tests confirmed viral infections. In intention-to-treat analysis, ILI was significantly higher among HCWs in the cloth masks group (RR=13.25 and 95% CI 1.74 to 100.97), compared with the medical masks group. The rate of ILI was also significantly higher in the cloth masks arm (RR=3.49 and 95% CI 1.00 to 12.17), compared with the control arm. Other outcomes were not statistically significant between the three arms.

Among the 68 laboratory-confirmed cases, 58 (85%) were due to rhinoviruses. Other viruses detected were hMPV (7 cases), influenza B (1 case), hMPV/rhinovirus co-infection (1 case) and influenza B/rhinovirus co-infection (1 case) (table 3). No influenza A or RSV infections were detected.

Compliance was significantly higher in the cloth mask arm (RR=2.41, 95% CI 2.01 to 2.88) and medical masks

Table 1 Demographic and other characteristics by arm of randomisation

Variable	Medical mask (% and 95% CI) (n=580)	Cloth mask (% and 95% CI) (n=569)	Control (% and 95% CI) (n=458)
Gender (male)	112/580	133/569	112/458
	19.3 (16.2 to 22.8)	23.4 (20.0 to 27.1)	24.5 (20.6 to 28.7)
Age (mean)	36 (35.6 to 37.3)	35 (34.6 to 36.3)	36 (35.1 to 37.0)
Education (postgraduate)	114/580	99/569	78/458
	19.7 (16.5 to 23.1)	17.4 (14.3 to 20.8)	17.0 (13.7 to 20.8)
Smoker (current/ex)	78/580	79/569	66/458
	13.4 (10.8 to 16.5)	13.9 (11.1 to 17.0)	14.4 (11.3 to 18.0)
Pre-existing illness*	66/580	70/569	47/458
	11.4 (9.0 to 14.2)	12.3 (9.8 to 15.3)	10.3 (7.8 to 13.4)
Influenza vaccination (yes)	21/580	21/569	15/458
	3.6 (2.4 to 5.4)	3.7 (2.4 to 5.6)	3.3 (2.0 to 5.3)
Staff (doctors)	176/580	165/569	134/458
	30.3 (26.6 to 34.3)	29.0 (25.3 to 32.9)	29.3 (25.1 to 33.7)
Number of hand washings per day (geometric mean)†	14 (13.8 to 15.4)	11 (10.9 to 11.9)	12 (11.5 to 12.7)
Number of patients had contact with (median and range)‡	21 (0 to 540)	21 (0 to 661)	18 (3 to 199)

*Includes asthma, immunocompromised and others.
†'Hand wash' variable was created by taking average of the number of hand washes performed by a healthcare worker (HCW) over the trial period. The variable was log transformed for the multivariate analysis.
‡'Number of patients had contact with' variable was created by taking average of the number of patients in contact with a HCW over the trial period. Median and range is presented in the table.

arm (RR=2.40, 95% CI 2.00 to 2.87), compared with the control arm. Figure 3 shows the percentage of participants who were compliant in the three arms. A post-hoc analysis adjusted for compliance and other potential confounders showed that the rate of ILI was significantly higher in the cloth mask arm (RR=13.00, 95% CI 1.69 to 100.07), compared with the medical masks arm (table 4). There was no significant difference between the medical mask and control arms. Hand washing was significantly protective against laboratory-confirmed viral infection (RR=0.66, 95% CI 0.44 to 0.97).

In the control arm, 170/458 (37%) used medical masks, 38/458 (8%) used cloth masks, and 245/458 (53%) used a combination of both medical and cloth masks during the study period. The remaining 1% either reported using a N95 respirator (n=3) or did not use any masks (n=2).

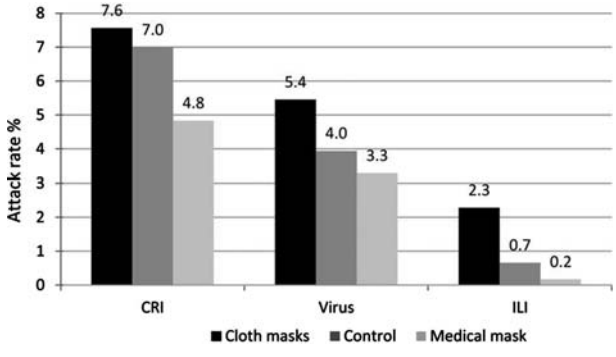


Figure 2 Outcomes in trial arms (CRI, clinical respiratory illness; ILI, influenza-like illness; Virus, laboratory-confirmed viruses).

Table 5 shows an additional analysis comparing all participants who used only a medical mask (from the control arm and the medical mask arm) with all participants who used only a cloth mask (from the control arm and the cloth arm). In the univariate analysis, all outcomes were significantly higher in the cloth mask group, compared with the medical masks group. After adjusting for other factors, ILI (RR=6.64, 95% CI 1.45 to 28.65) and laboratory-confirmed virus (RR=1.72, 95% CI 1.01 to 2.94) remained significantly higher in the cloth masks group compared with the medical masks group.

Table 6 compares the outcomes in the medical mask arm with two previously published trials.[8 9] This shows that while the rates of CRI were significantly higher in one of the previously published trials, the rates of laboratory-confirmed viruses were not significantly different between the three trials for medical mask use.

On average, HCWs worked for 25 days during the trial period and washed their cloth masks for 23/25 (92%) days. The most common approach to washing cloth masks was self-washing (456/569, 80%), followed by combined self-washing and hospital laundry (91/569, 16%), and only hospital laundry (22/569, 4%). Adverse events associated with facemask use were reported in 40.4% (227/562) of HCWs in the medical mask arm and 42.6% (242/568) in the cloth mask arm (p value 0.450). General discomfort (35.1%, 397/1130) and breathing problems (18.3%, 207/1130) were the most frequently reported adverse events.

Table 2 Intention-to-treat analysis

	CRI N (%)	RR (95% CI)	ILI N (%)	RR (95% CI)	Laboratory-confirmed viruses N (%)	RR (95% CI)
Medical mask*	28/580 (4.83)	Ref	1/580 (0.17)	Ref	19/580 (3.28)	Ref
Cloth masks†	43/569 (7.56)	1.57 (0.99 to 2.48)	13/569 (2.28)	**13.25 (1.74 to 100.97)**	31/569 (5.45)	1.66 (0.95 to 2.91)
Control‡	32/458 (6.99)	1.45 (0.88 to 2.37)	3/458 (0.66)	3.80 (0.40 to 36.40)	18/458 (3.94)	1.20 (0.64 to 2.26)

Bold typeface indicates statistically significant.
*p Value from cluster adjusted χ^2 tests is 0.510 and intracluster correlation coefficients is 0.065.
†p Value from cluster adjusted χ^2 tests is 0.028 and intracluster correlation coefficients is 0.029.
‡p Value from cluster adjusted χ^2 tests is 0.561 and intracluster correlation coefficients is 0.068.
CRI, clinical respiratory illness; ILI, influenza-like illness; RR, relative risk.

Laboratory tests showed the penetration of particles through the cloth masks to be very high (97%) compared with medical masks (44%) (used in trial) and 3M 9320 N95 (<0.01%), 3M Vflex 9105 N95 (0.1%).

DISCUSSION

We have provided the first clinical efficacy data of cloth masks, which suggest HCWs should not use cloth masks as protection against respiratory infection. Cloth masks resulted in significantly higher rates of infection than medical masks, and also performed worse than the control arm. The controls were HCWs who observed standard practice, which involved mask use in the majority, albeit with lower compliance than in the intervention arms. The control HCWs also used medical masks more often than cloth masks. When we analysed all mask-wearers including controls, the higher risk of cloth masks was seen for laboratory-confirmed respiratory viral infection.

The trend for all outcomes showed the lowest rates of infection in the medical mask group and the highest rates in the cloth mask arm. The study design does not allow us to determine whether medical masks had efficacy or whether cloth masks were detrimental to HCWs by causing an increase in infection risk. Either possibility, or a combination of both effects, could explain our results. It is also unknown whether the rates of infection observed in the cloth mask arm are the same or higher than in HCWs who do not wear a mask, as almost all participants in the control arm used a mask. The physical properties of a cloth mask, reuse, the frequency and effectiveness of cleaning, and increased moisture retention, may potentially increase the infection risk for HCWs. The virus may survive on the surface of the facemasks,[29] and modelling studies have quantified the contamination levels of masks.[30] Self-contamination through repeated use and improper doffing is possible. For example, a contaminated cloth mask may transfer pathogen from the mask to the bare hands of the wearer. We also showed that filtration was extremely poor (almost 0%) for the cloth masks. Observations during SARS suggested double-masking and other practices increased the risk of infection because of moisture, liquid diffusion and pathogen retention.[31] These effects may be associated with cloth masks.

We have previously shown that N95 respirators provide superior efficacy to medical masks,[8 9] but need to be worn continuously in high-risk settings to protect HCWs.[9] Although efficacy for medical masks was not shown, efficacy of a magnitude that was too small to be detected is possible.[8 9] The magnitude of difference between cloth masks and medical masks in the current study, if explained by efficacy of medical masks alone, translates to an efficacy of 92% against ILI, which is possible, but not consistent with the lack of efficacy in the two previous RCTs.[8 9] Further, we found no significant difference in rates of virus isolation in medical mask users between the three trials, suggesting that the results of this study could be interpreted as partly being explained by a detrimental effect of cloth masks. This is further supported by the fact that the rate of virus isolation in the no-mask control group in the first Chinese RCT was 3.1%, which was not significantly different to the rates of virus isolation in the medical mask arms in any of the three trials including this one. Unlike the previous RCTs, circulating influenza and RSV were almost completely absent during this study,

Table 3 Type of virus isolated

Study arm	hMPV	Rhino	Influenza B virus	hMPV & rhino	Influenza B virus & rhino	Total
Medical masks arm	1	16	1	1	0	19
Cloth mask arm	4	26	0	0	1	31
Control arm	2	16	0	0	0	18
Total	7	58	1	1	1	68

hMPV, human metapneumovirus; Rhino, rhinoviruses.

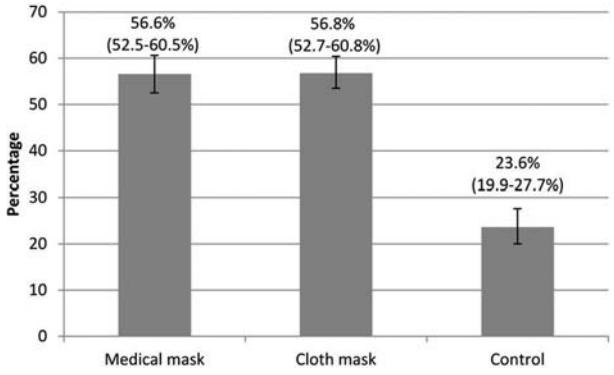


Figure 3 Compliance with the mask wearing—mask wearing more than 70% of working hours.

with rhinoviruses comprising 85% of isolated pathogens, which means the measured efficacy is against a different range of circulating respiratory pathogens. Influenza and RSV predominantly transmit through droplet and contact routes, while Rhinovirus transmits through multiple routes, including airborne and droplet routes.[32 33] The data also show that the clinical case definition of ILI is non-specific, and captures a range of pathogens other than influenza. The study suggests medical masks may be protective, but the magnitude of difference raises the possibility that cloth masks cause an increase in infection risk in HCWs. Further, the filtration of the medical mask used in this trial was poor, making extremely high efficacy of medical masks unlikely, particularly given the predominant pathogen was rhinovirus, which spreads by the airborne route. Given the obligations to HCW occupational health and safety, it is important to consider the potential risk of using cloth masks.

In many parts of the world, cloth masks and medical masks may be the only options available for HCWs. Cloth masks have been used in West Africa during the Ebola outbreak in 2014, due to shortages of PPE, (personal communication, M Jalloh). The use of cloth masks is recommended by some health organisations, with caveats.[34–36] In light of our study, and the obligation to ensure occupational health and safety of HCWs, cloth masks should not be recommended for HCWs, particularly during AGPs and in high-risk settings such as emergency, infectious/respiratory disease and intensive care wards. Infection control guidelines need to acknowledge the widespread real-world practice of cloth masks and should comprehensively address their use. In addition, other important infection control measure such as hand hygiene should not be compromised. We confirmed the protective effects of hand hygiene against laboratory-confirmed viral infection in this study, but mask type was an independent predictor of clinical illness, even adjusted for hand hygiene.

A limitation of this study is that we did not measure compliance with hand hygiene, and the results reflect self-reported compliance, which may be subject to recall or other types of bias. Another limitation of this study is the lack of a no-mask control group and the high use of masks in the controls, which makes interpretation of the results more difficult. In addition, the quality of paper and cloth masks varies widely around the world, so the results may not be generalisable to all settings. The lack of influenza and RSV (or asymptomatic infections) during the study is also a limitation, although the predominance of rhinovirus is informative about pathogens transmitted by the droplet and airborne routes in this setting. As in previous studies, exposure to infection outside the workplace could not be estimated, but we would assume it to be equally distributed between trial arms. The major strength of the randomised trial study design is in ensuring equal distribution of confounders and effect modifiers (such as exposure outside the workplace) between trial arms.

Cloth masks are used in resource-poor settings because of the reduced cost of a reusable option. Various types of cloth masks (made of cotton, gauze and other fibres) have been tested in vitro in the past and show lower filtration capacity compared with disposable masks.[7] The protection afforded by gauze masks increases with the fineness of the cloth and the number of layers,[37] indicating potential to develop a more effective cloth mask, for example, with finer weave, more layers and a better fit.

Cloth masks are generally retained long term and reused multiple times, with a variety of cleaning methods and widely different intervals of cleaning.[34] Further studies are required to determine if variations in frequency and type of cleaning affect the efficacy of cloth masks.

Table 4 Multivariable cluster-adjusted log-binomial model to calculate RR for study outcomes

	CRI RR (95% CI)	ILI RR (95% CI)	Laboratory-confirmed viruses RR (95% CI)
Medical masks arm	Ref	Ref	Ref
Cloth mask arm	1.56 (0.97 to 2.48)	**13.00 (1.69 to 100.07)**	1.54 (0.88 to 2.70)
Control arm	1.51 (0.90 to 2.52)	4.64 (0.47 to 45.97)	1.09 (0.57 to 2.09)
Male	0.67 (0.41 to 1.12)	1.03 (0.34 to 3.13)	0.65 (0.34 to 1.22)
Vaccination	0.83 (0.27 to 2.52)	1.74 (0.24 to 12.56)	1.27 (0.41 to 3.92)
Hand washing	0.91 (0.66 to 1.26)	0.94 (0.40 to 2.20)	**0.66 (0.44 to 0.97)**
Compliance	1.14 (0.77 to 1.69)	1.86 (0.67 to 5.21)	0.86 (0.53 to 1.40)

Bold typeface indicates statistically significant.
CRI, clinical respiratory illness; ILI, influenza-like illness; RR, relative risk.

Table 5 Univariate and adjusted analysis comparing participants who used medical masks and cloth masks*

	Univariate RR (95% CI)	Adjusted RR (95% CI)
CRI		
Medical mask (35/750, 4.67%)	Ref	Ref
Cloth mask (46/607, 7.58%)	**1.62 (1.06 to 2.49)**	1.51 (0.97 to 2.32)
Male	0.60 (0.32 to 1.12)	0.58 (0.31 to 1.08)
Vaccination	0.66 (0.17 to 2.62)	0.68 (0.17 to 2.67)
Hand washing	0.81 (0.58 to 1.15)	0.84 (0.59 to 1.20)
Compliance	1.01 (1.00 to 1.03)	1.01 (1.00 to 1.02)
ILI		
Medical mask (2/750, 0.27%)	Ref	Ref
Cloth mask (13/607, 2.14%)	**8.03 (1.82 to 35.45)**	**6.64 (1.45 to 28.65)**
Male	0.95 (0.27 to 3.35)	0.92 (0.26 to 3.22)
Vaccination	1.87 (0.25 to 13.92)	1.97 (0.27 to 14.45)
Hand washing	0.56 (0.24 to 1.27)	0.61 (0.23 to 1.57)
Compliance	1.04 (1.01 to 1.08)	1.04 (1.00 to 1.08)
Laboratory-confirmed viruses		
Medical mask (22/750, 2.93%)	Ref	Ref
Cloth mask (34/607, 5.60%)	**1.91 (1.13 to 3.23)**	**1.72 (1.01 to 2.94)**
Male	0.64 (0.30 to 1.33)	0.61 (0.29 to 1.27)
Vaccination	0.97 (0.24 to 3.86)	1.03 (0.26 to 4.08)
Hand washing	0.61 (0.41 to 0.93)	0.65 (0.42 to 1.00)
Compliance	1.00 (0.99 to 1.02)	1.0 (0.99 to 1.02)

Bold typeface indicates statistically significant.
*The majority (456/458) of HCWs in the control arm used a mask. Controls who exclusively used a medical mask were categorised and analysed with the medical mask arm participants; and controls who exclusively wore a cloth mask were categorised and analysed with the cloth mask arm.
CRI, clinical respiratory illness; HCWs, healthcare workers; ILI, influenza-like illness; RR, relative risk.

Table 6 A comparison of outcome data for the medical mask arm with medical mask outcomes in previously published RCTs

	CRI N (%)	RR (95% CI)	ILI N (%)	RR (95% CI)	Laboratory-confirmed viruses N (%)	RR (95% CI)
Vietnam trial	28/580 (4.83)	Ref	1/580 (0.17)	Ref	19/580 (3.28)	Ref
Published RCT China 1[8]	33/492 (6.70)	1.40 (0.85 to 2.26)	3/492 (0.61)	3.53 (0.37 to 33.89)	13/492 (2.64)	0.80 (0.40 to 1.62)
Published RCT China 2[9]	98/572 (17.13)	**3.54 (2.37 to 5.31)**	4/572 (0.70)	4.06 (0.45 to 36.18)	19/572 (3.32)	1.01 (0.54 to 1.89)

Bold typeface indicates statistically significant.
CRI, Clinical respiratory illness; ILI, influenza-like illness; RCT, randomised clinical trial; RR, relative risk.

Pandemics and emerging infections are more likely to arise in low-income or middle-income settings than in wealthy countries. In the interests of global public health, adequate attention should be paid to cloth mask use in such settings. The data from this study provide some reassurance about medical masks, and are the first data to show potential clinical efficacy of medical masks. Medical masks are used to provide protection against droplet spread, splash and spray of blood and body fluids. Medical masks or respirators are recommended by different organisations to prevent transmission of Ebola virus, yet shortages of PPE may result in HCWs being forced to use cloth masks.[38–40] In the interest of providing safe, low-cost options in low income countries, there is scope for research into more effectively designed cloth masks, but until such research is carried out, cloth masks should not be recommended. We also recommend that infection control guidelines be updated about cloth mask use to protect the occupational health and safety of HCWs.

Author affiliations
[1]Faculty of Medicine, School of Public Health and Community Medicine, University of New South Wales, Sydney, Australia
[2]National Institute of Hygiene and Epidemiology, Hanoi, Vietnam
[3]Institute for Clinical Pathology and Medical Research, Westmead Hospital and University of Sydney, Sydney, New South Wales, Australia
[4]Beijing Centers for Disease Control and Prevention, Beijing, China

Acknowledgements The authors would like to thank the staff members from the National Institute of Hygiene and Epidemiology, Hanoi, Vietnam, who were involved with the trial. They thank as well to the staff from the Hanoi hospitals who participated. They also acknowledge the support of 3M for testing of filtration of the facemasks. 3M was industry partner in the ARC linkage project

grant; however they were not involved in study design, data collection or analysis. The 3M products were not used in this study.

Contributors CRM was the lead investigator, and responsible for the conception and design of the trial, obtaining the grant funding, overseeing the whole study, analysing the data and writing of the report. HS contributed to overseeing the study, staff training, form/database development and drafting of the manuscript. TCD was responsible for overseeing the study, database management, recruitment, training and revision of the manuscript. NTH was responsible for the implementation of research and revision of the manuscript. PTN was responsible for the laboratory testing in Vietnam. AAC contributed to the statistical analysis and drafting of the manuscript. BR was responsible for the statistical analysis and revision of the manuscript. DED contributed to the laboratory technical assistance and revision of the manuscript. QW assisted in comparing the rates of infection from two previous RCTs conducted in China and revision of the manuscript.

Funding Funding to conduct this study was received from the Australian Research Council (ARC) (grant number LP0990749).

Competing interests CRM has held an Australian Research Council Linkage Grant with 3M as the industry partner, for investigator-driven research. 3M has also contributed masks and respirators for investigator-driven clinical trials. CRM has received research grants and laboratory testing as in-kind support from Pfizer, GSK and Bio-CSL for investigator-driven research. HS had a NHMRC Australian-based Public Health Training Fellowship at the time of the study (1012631). She has also received funding from vaccine manufacturers GSK, bio-CSL and Sanofi Pasteur for investigator-driven research and presentations. AAC used filtration testing of masks for his PhD thesis conducted by 3M Australia.

Ethics approval National Institute for Hygiene and Epidemiology (NIHE) (approval number 05 IRB) and the Human Research Ethics Committee of the University of New South Wales (UNSW), Australia, (HREC approval number 10306).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. World Health Organization (WHO). Global Alert and Response (GAR), Pandemic (H1N1) 2009—update 76 (cited 27 Apr 2012). http://www.who.int/csr/don/2009_11_27a/en/index.html
2. World Health Organization (WHO). Human infection with avian influenza A(H7N9) virus—update (cited 8 May 2013). http://www.who.int/csr/don/2013_05_07/en/index.html
3. Bermingham A, Chand MA, Brown CS, *et al*. Severe respiratory illness caused by a novel coronavirus, in a patient transferred to the United Kingdom from the Middle East, September 2012. *Euro Surveill* 2012;17:20290.
4. Pollack MP, Pringle C, Madoff LC, *et al*. Latest outbreak news from ProMED-mail: novel coronavirus—Middle East. *Int J Infect Dis* 2013;17:e143–4.
5. World Health Organization (WHO). Global Alert and Response (GAR). Ebola virus disease update—west Africa 2014 (cited 28 Aug 2014). http://www.who.int/csr/don/2014_08_28_ebola/en/
6. Chughtai AA, Seale H, MacIntyre CR. Use of cloth masks in the practice of infection control—evidence and policy gaps. *Int J Infect Control* 2013;9:1–12.
7. Quesnel LB. The efficiency of surgical masks of varying design and composition. *Br J Surg* 1975;62:936–40.
8. MacIntyre CR, Wang Q, Cauchemez S, *et al*. A cluster randomized clinical trial comparing fit-tested and non-fit-tested N95 respirators to medical masks to prevent respiratory virus infection in health care workers. *Influenza Other Respir Viruses* 2011;5:170–9.
9. MacIntyre CR, Wang Q, Seale H, *et al*. A randomised clinical trial of three options for N95 respirators and medical masks in health workers. *Am J Respir Crit Care Med* 2013;187:960–6.
10. Chughtai AA, MacIntyre CR, Zheng Y, *et al*. Examining the policies and guidelines around the use of masks and respirators by healthcare workers in China, Pakistan and Vietnam. *J Infect Prev* 2015;16:68–74.
11. Chughtai AA, Seale H, Chi Dung T, *et al*. Current practices and barriers to the use of facemasks and respirators among hospital-based health care workers in Vietnam. *Am J Infect Control* 2015;43:72–7.
12. Pang X, Zhu Z, Xu F, *et al*. Evaluation of control measures implemented in the severe acute respiratory syndrome outbreak in Beijing. *JAMA* 2003;290:3215–21.
13. Yang P, Seale H, MacIntyre C, *et al*. Mask-wearing and respiratory infection in healthcare workers in Beijing, China. *Braz J Infect Dis* 2011;15:102–8.
14. Horton R. Medical journals: evidence of bias against the diseases of poverty. *Lancet* 2003;361:712–13.
15. MacIntyre CR, Chughtai AA. Facemasks for the prevention of infection in healthcare and community settings. *BMJ* 2015;350:h694.
16. Chughtai AA, Seale H, MacIntyre CR. Availability, consistency and evidence-base of policies and guidelines on the use of mask and respirator to protect hospital health care workers: a global analysis. *BMC Res Notes* 2013;6:1–9.
17. MacIntyre C, Cauchemez S, Dwyer D, *et al*. Face mask use and control of respiratory virus transmission in households. *Emerg Infect Dis* 2009;15:233–41.
18. Buecher C, Mardy S, Wang W, *et al*. Use of a multiplex PCR/RT-PCR approach to assess the viral causes of influenza-like illnesses in Cambodia during three consecutive dry seasons. *J Med Virol* 2010;82:1762-72 [Epub ahead of print 1 Sep 2010].
19. Higuchi R, Fockler C, Dollinger G, *et al*. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology* (*N Y*) 1993;11:1026–30 [Epub ahead of print 1 Sep 2010].
20. Hummel KB, Lowe L, Bellini WJ, *et al*. Development of quantitative gene-specific real-time RT-PCR assays for the detection of measles virus in clinical specimens. *J Virol Methods* 2006;132:166–73 [Epub ahead of print 11 Sep 2005].
21. Mackay IM. *Real-time PCR in microbiology*. Caister Academic Press, 2007.
22. Wang W, Cavailler P, Ren P, *et al*. Molecular monitoring of causative viruses in child acute respiratory infection in endemo-epidemic situations in Shanghai. *J Clin Virol* (PASCV) 2010;49:211–8.
23. Thi TN, Deback C, Malet I, *et al*. Rapid determination of antiviral drug susceptibility of herpes simplex virus types 1 and 2 by real-time PCR. *Antiviral Res* 2006;69:152–7.
24. Standards Australia Limited/Standards New Zealand. Respiratory protective devices. Australian/New Zealand Standard. AS/NZS 1716: 2012.
25. Donner A, Klar N. *Design and analysis of cluster randomization trials in health research*. London: Oxford University Press Inc, 2000.
26. Campbell MK, Elbourne DR, Altman DG, *et al*. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004;328:702–8.
27. Vittinghoff E, Glidden DV, Shiboski SC, *et al*. *Regression methods in biostatistics*. 2nd edn. New York: Springer-Verlag, 2012.
28. StataCorp. *Stata 12 base reference manual*. College Station, TX: Stata Press, 2011.
29. Osterholm MT, Moore KA, Kelley NS, *et al*. Transmission of Ebola viruses: what we know and what we do not know. *mBio* 2015;6: e00137–15.
30. Fisher EM, Noti JD, Lindsley WG, *et al*. Validation and application of models to predict facemask influenza contamination in healthcare settings. *Risk Anal* 2014;34:1423–34.
31. Li Y, Wong T, Chung J, *et al*. In vivo protective performance of N95 respirator and surgical facemask. *Am J Ind Med* 2006;49:1056–65.
32. Dick EC, Jennings LC, Mink KA, *et al*. Aerosol transmission of rhinovirus colds. *J Infect Dis* 1987;156:442–8.
33. Bischoff WE. Transmission route of rhinovirus type 39 in a monodispersed airborne aerosol. *Infect Control Hosp Epidemiol* 2010;31:857–9.
34. Institute of Medicine (IOM). Reusability of Facemasks During an Influenza Pandemic: Facing the Flu—Committee on the Development of Reusable Facemasks for Use During an Influenza Pandemic. National Academy of Sciences, 2006.
35. Center for Disease Control and Prevention and World Health Organization. *Infection control for viral haemorrhagic fevers in the African health care setting*. Atlanta: Centers for Disease Control and Prevention, 1998:1–198.

36. World Health Organization (WHO). Guidelines for the prevention of tuberculosis in health care facilities in resource limited settings, 1999.
37. Weaver GH. Droplet infection and its prevention by the face mask. *J Infect Dis* 1919;24:218–30.
38. MacIntyre CR, Chughtai AA, Seale H, *et al*. Respiratory protection for healthcare workers treating Ebola virus disease (EVD): are facemasks sufficient to meet occupational health and safety obligations? *Int J Nurs Stud* 2014;51:1421–6.
39. Center for Disease Control and Prevention (CDC). Guidance on Personal Protective Equipment To Be Used by Healthcare Workers During Management of Patients with Ebola Virus Disease in U.S. Hospitals, Including Procedures for Putting On (Donning) and Removing (Doffing). 2014 (cited 23 Oct 2014). http://www.cdc.gov/vhf/ebola/hcp/procedures-for-ppe.html
40. World Health Organiszation (WHO). Infection prevention and control guidance for care of patients in health-care settings, with focus on Ebola. 2014 (cited 23 Oct 2014). http://www.who.int/csr/resources/publications/ebola/filovirus_infection_control/en/

Review

The Royal Society of Medicine

Journal of the Royal Society of Medicine; 2015, Vol. 108(6) 223–228
DOI: 10.1177/0141076815583167

Unmasking the surgeons: the evidence base behind the use of facemasks in surgery

Charlie Da Zhou[1], Pamela Sivathondan[2] and Ashok Handa[2]
[1]New College, University of Oxford, Oxford OX1 3BN, UK
[2]Nuffield Department of Surgical Sciences, University of Oxford, Oxford OX3 9DU, UK
Corresponding author: Charlie Da Zhou. Email: Charlie.zhou@new.ox.ac.uk

Summary

The use of surgical facemasks is ubiquitous in surgical practice. Facemasks have long been thought to confer protection to the patient from wound infection and contamination from the operating surgeon and other members of the surgical staff. More recently, protection of the theatre staff from patient-derived blood/bodily fluid splashes has also been offered as a reason for their continued use. In light of current NHS budget constraints and cost-cutting strategies, we examined the evidence base behind the use of surgical facemasks.

Examination of the literature revealed much of the published work on the matter to be quite dated and often studies had poorly elucidated methodologies. As a result, we recommend caution in extrapolating their findings to contemporary surgical practice. However, overall there is a lack of substantial evidence to support claims that facemasks protect either patient or surgeon from infectious contamination. More rigorous contemporary research is needed to make a definitive comment on the effectiveness of surgical facemasks.

Keyword

Surgery

Picture a surgeon operating in a theatre, and chances are that you will imagine them wearing a surgical facemask. Masks are a quintessential part of the surgical attire that has become so deeply ingrained in the public perception of the profession. However, even today, it remains unclear as to whether they confer any tangible benefits to surgical outcomes. As 'efficiency' and 'cost-cutting' have increasingly become the *topics du jour* in the National Health Service, it seems reasonable to assess the efficacy, effectiveness and cost-to-benefit ratio for this particular component of the surgical uniform.

Methodology

We searched the PubMed journal database and Google Scholar with the search terms 'surgical facemask/mask', 'splash', 'contamination', 'infection' and 'outcomes' in order to identify salient publications. We also searched the guidance on surgical site infection from the National Institute for Health and Care Excellence. Furthermore, a manual search of reference lists from relevant papers was performed.

Contemporary attitudes to the surgical mask

A contemporary questionnaire-based study, which attempted to assess the attitudes of surgeons, revealed that 96% of responders wore facemasks.[1] About equal numbers did so with the primary aim of protecting the patients compared to protecting themselves. However, it was also found that 20% of responding surgeons wore the mask for the sole purpose of respecting tradition. Furthermore, 30% of responding surgeons felt that masks could make surgery more difficult by increasing breath condensation on spectacles, endoscopes and microscopes and thereby obscuring vision.

In May 2014, the first installation of the Glass Surgery project was broadcast to viewers around the world. This project, based at the Barts and the London School of Medicine and Dentistry, was the first of its kind to live-stream a surgical procedure, using new Google Glass technology, to any medical student or trainee with an internet connection. Mr Ahmed, the lead colorectal surgeon, elected not to wear a mask while performing the open right hemicolectomy and partial liver resection in question. In the immediate aftermath of the broadcast, Mr Ahmed came under scrutiny from various medical comment threads, blogs and chat rooms on the Internet questioning his decision to omit the facemask and whether this might have compromised patient safety.

Protection of the patient

The facemask has been used in surgical settings for over a hundred years;[2] first described in 1897, at its

inception, it consisted merely of a single layer of gauze to cover the mouth,[3] and its primary function was to protect the patient from contamination and surgical site infection. This practice was substantiated, at the time, by a recent discovery which demonstrated that bacteria could be disseminated from the nose and mouth during normal conversation as observed by bacterial colony growth on strategically placed agar plates in theatres. In the 1940s and 1950s, antibiotics and aseptic technique came to the forefront of infection control strategies within the surgical setting. Until recently, it has remained unclear as to whether bacterial colony growth on an agar plate was a direct correlate of surgical site infections and also whether the purpose of the surgical mask has been superseded by more modern strategies of infection control.

In order to advocate the validity of an intervention in medicine, it must satisfy three levels of evidence: efficacy, effectiveness and cost-effectiveness.[4] In the context of facemask, efficacy is whether masks prevent the propagation of droplets derived from the mouth and nose of the operating staff. Effectiveness is whether efficacy translates into a significant reduction in surgical site infection morbidity and mortality. And finally, cost-effectiveness determines whether the cost-to-benefit ratio of this effect would be desirable compared to an alternative course of action.

Intuition would suggest that facemasks offer a physical barrier preventing the emanation of droplets from the oral or nasal passages and therefore satisfy the efficacy requirement of the evidence ladder. However, there are a number of different hypotheses as to why this may not be the case. 'Venting' is a phenomenon whereby air leaks at the interface between mask and face which can act to disperse potential contaminants originating from the pharynx.[5] The accumulation of moisture, during prolonged usage, may exacerbate this problem by increasing resistance to air flow through the filter itself. Moisture accumulation is also thought to facilitate the movement of contaminants through the material of the mask itself by capillary action. These bacteria can subsequently be dislodged by movement. Friction at the face/mask interface has also been demonstrated to disperse skin scales which can further contribute towards wound contamination.[6]

In the modern era, there has also been a scarcity of experimental evidence to support the effectiveness of facemasks in the prevention of surgical site infections. The earliest retrospective studies[7] failed to demonstrate any statistically significant improvement in surgical site infection rates following the use of masks. Indeed, the latest National Institute for Health and Care Excellence guidelines on the matter do not require operating staff to wear a mask in theatre.[8] This decision was based primarily upon the findings of a Cochrane systematic review.[9] This review was guided by the findings of two particular randomised/quasi-randomised control trials.[10,11] The latest update of this review,[12] which was amended after the publication of current National Institute for Health and Care Excellence guidelines, included one further study.[13]

The Cochrane review[12] searched through six established databases (Appendix 1) looking for randomised control trials and quasi-randomised control trials investigating surgical outcomes comparing the use of disposable surgical masks with the use of no masks. The authors limited the scope of their analysis only to patients undergoing clean procedures (whereby the operating procedure does not enter a body cavity or viscus normally colonised by bacteria). The review chose not to investigate the role of mask in clean-contaminated, contaminated or dirty wounds as one would expect that masks would contribute less towards the prevention of surgical site infections under such circumstances. Primary outcomes of postoperative surgical wound infection and secondary outcomes of costs, length of hospital stay and mortality rates were ascertained.

Three studies were identified as fulfilling all the selection criteria of the review.[10,11,13] A total of 2106 participants were identified across the three studies (Table 1). All the studies reported on the primary outcome of postoperative surgical wound infection, none of the studies reported on any of the secondary outcomes. Furthermore, identified studies were assessed for risk of bias based on eight specific criteria (Table 2).

Statistical analysis of the extracted data revealed no statistically significant association between mask usage and the incidence of surgical site infection. The study concluded that 'it is unclear whether the wearing of surgical facemasks by members of the surgical team has any impact on surgical wound infection rates for patients undergoing clean surgery'. However, each of the studies included could be criticised for risk of bias (Table 2). Indeed, the Webster study, arguably the most rigorous of the three, only investigated the impact of mask on non-scrubbed members of the surgical team. There is uncertainty over whether the findings of some of these studies are applicable to contemporary surgical practice.

Based upon the findings of this review, National Institute for Health and Care Excellence guidelines state that there is 'limited evidence concerning the use of non-sterile theatre wear' such as surgical masks when trying to minimise the risk of surgical site infection, although there was an overall 'consensus that wearing non-sterile theatre wear is important in maintaining theatre discipline'. This latter

Table 1. Characteristics of included studies.[12]

Study	Methods	Participants	Outcomes	Notes
Chamberlain and Houang[10]	Quasi-randomised controlled trial	41 female patients undergoing gynaecology surgery. 24 clean and 17 non-clean. Of the clean surgeries: masked cohort $n = 14$, unmasked cohort $n = 10$	Wound infection defined as serious enough to warrant antibiotics in 2 cases and via high vaginal swab in third case. Follow-up until discharge only. No postoperative wound infections in the masked group and 3/10 (30%) in the non-masked group (no statistically significant difference: OR 0.07, 95% CI 0.00–1.63)	Study discontinued due to 3 surgical wound infections in unmasked group, although not proven as causal. Data extracted for clean surgery only. Unit of analysis error present.
Tunevall[11]	Quasi-randomised controlled trial	3088 patients undergoing general, vascular, breast, acute and elective surgery. 1429 clean and 1659 unclean. Of the clean surgeries: masked cohort $n = 706$, unmasked cohort $n = 723$	Wound infection defined as visible pus and/or cellulitis without pus requiring debridement, drainage and/or antibiotics. Duration of follow-up not stated but until after discharge from ward. 13/706 (1.8%) postoperative wound infections in the masked group and 10/723 (1.4%) in the non-masked group (no statistically significant difference: OR 1.34, 95% CI 0.58–3.07)	Data extracted from clean surgery only. Patients had 2 to 3 body washes preoperatively with 4% chlorhexidine prior to elective surgery. In most acute cases, at least one body wash was given. Unit of analysis error present.
Webster et al.[13]	Randomised controlled trial	811 patients undergoing gynaecological, obstetric, general (open), general (laparoscopic), urology and breast surgery. 660 clean and 151 non-clean. Of the clean surgeries: masked cohort $n = 313$, unmasked cohort $n = 340$	Wound infection defined by criterial used by National Nosocomial Infection Surveillance System of Australia. Clean surgery masked cohort, mean follow-up 33.4 days (SD 22.1). Clean surgery unmasked cohort, mean follow-up 33.4 days (SD 22.8). Infection rate 33/313 (10.5%) in the masked group and 31/340 (9.1%) in the non-masked group (no statistically significant difference: OR 1.17, 95% CI 0.70–1.97)	Scrubbed staff were not included in trial. Data extracted from clean surgery only. Missing data for 7 clean cases. Unit of analysis error present.

statement seems to be a rather vague and likely unfounded assertion which implies a correlation between dress code, staff discipline and thereby patient safety outcomes. This may reflect a reluctance among the medical profession to deviate from embedded tradition as reflected in Leyland and McCloy's questionnaire study.[1] Alternatively, it may reflect a prevailing intuition that surgical masks ought to protect against surgical site infections.

Table 2. Assessment for risk of bias in included studies.[12]

Study	1	2	3	4	5	6	7	8	9
Chamberlain and Houang[10]	?	?	?	L	L	L	?	H	?
Tunevall[11]	H	H	?	H	L	L	?	L	L
Webster et al.[13]	L	L	L	L	?	L	L	L	L

L: low risk; ?: uncertain risk; H: high risk.
Bias was assessed by the following aspects: (1) method of randomisation: how the randomisation schedule was generated, the method of randomisation, e.g. envelopes, computer etc., (2) allocation concealment, (3) blinding of patients (recipients), (4) blinding of outcome assessors to wearing of masks, (5) extent of loss to follow-up and use of intention-to-treat analysis, (6) source of funding, (7) selective reporting, (8) early stopping and (9) baseline comparability of treatment and control groups.

Unfortunately, publically available information regarding the financial costs of facemask usage on the National Health Service is lacking. However, as part of the Freedom of Information Publication Scheme, the data are available for the West Hertfordshire Hospitals NHS Trust which purchased 44,482 single-use facemasks in 2012.[14] During this year, the West Hertfordshire Hospitals NHS Trust performed a total of 63,250 operative procedures or interventions.[15] Extrapolation to the 10,594,814 total operative procedures and interventions carried out across NHS England during the same period[15] would equate to an annual procurement of almost 7.5 million single-use masks across hospitals in England. The NHS Atlas of Procurement lists the per unit expenditure of surgical facemasks to be between £0.34 to £1.22, depending on trust and supplier.[16] This suggests that annual NHS England expenditure on facemasks lies somewhere in the region of £2.5 to £9.1 million.

Hospital-acquired infections, of which surgical site infections are a subset, are a major problem for all health systems. Media coverage, in recent times, has heightened public awareness of their associated morbidity and mortality. Their socioeconomic impact is also substantial,[17] and it is estimated that iatrogenic infection increases the duration of average hospital stay by a factor of 2.5 while incurring almost three times the monetary cost of uninfected patients. Across the whole of the United Kingdom, it is estimated that annually hospital-acquired infections cost the National Health Service almost £1 billion in excess expenditure and a loss of 3.6 million bed days. Personal costs for the patients are also affected as their return to normal daily activity and employment are delayed.

Given the uncertainty in effectiveness of facemasks in preventing surgical site infection, it is impossible to perform a cost-to-benefit analysis on mask usage. It is clear, however, that the National Health Service expenditure on facemasks is a mere fraction of the costs incurred due to hospital-acquired infections.

Protection of the surgeon

An increasingly prevalent belief, in favour of mask usage, is the idea that they also confer some degree of protection to the operating staff from patient-derived infectious material.[18] Most obviously, they can act as a physical barrier against blood and bodily fluid splashes during surgery. One prospective study revealed that facemasks prevented blood/bodily fluid splashes that would have otherwise contaminated the surgeon's face in 24% of procedures.[19] The incidence of blood/bodily fluid splashes varies substantially between settings and between individuals. The risk is modified by the role of surgical staff (lead surgeons are at higher risk than first assistants, who in turn have a higher risk than scrub nurses), by surgical specialty as well as by surgical technique.[19,20] The frequency of blood/bodily fluid splashed has been reported to be as high as 62.5% in lead surgeons performing Caesarean section.[20]

Despite clear evidence that facemasks act to protect the theatre staff from macroscopic facial contamination, there are studies to suggest that they fail to protect surgeons from potentially hazardous submicrometre contaminants.[21] This corresponds roughly to the size range of infectious bacteria while viruses are even smaller. Therefore, the protection that masks confer in the form of macroscopic facial contamination may not necessarily extend towards any microscopic infectious agents present within that contamination.

Proponents of the surgical facemask may argue that even if they fail to completely negate the risks of infection they are likely to reduce exposure in a dose-dependent manner. While this field has not been extensively investigated, preliminary work suggests that facemasks fail to confer any degree of protection from infection due to streptococcal and staphylococcal bacterial species[22] or hepatitis B virus.[23] Furthermore, a facemask splash may promote a false sense of security, as surgeons may be less likely to report these as an

occupational exposure to bodily fluid compared to frank facial contamination.

Tying things together

In surgery, there are many aspects of current clinical practice that do not necessarily have an established evidence base. Indeed, it is permissible to bypass the evidence ladder when an intervention is so convincing that it is possible to discern its effect signal from noise by observation alone.[24] In such circumstances, interventions have a very clear mechanistic cause and effect relationship. Historically, it may have been thought that surgical masks fulfilled such criteria. This would explain why published literature examining surgical mask effectiveness has been lacking despite their ubiquitous nature within the surgical profession.

What literature that is available on the subject tends to be dated with poorly explained methodology. There is also uncertainty over whether the results of such studies can be extrapolated to current surgical practice given the advent of new antiseptic techniques since they were completed. The evidence base investigating the effects of facemask usage on patient-based outcomes is, in general, more extensive than that of surgeon-centred outcomes. Facemasks do have a clear role in maintaining the social cleanliness of surgical staff, but evidence is lacking to suggest that they confer protection from infection either to patients or to the surgeons that wear them.

Given that there is no evidence that they cause any harm either, proponents would rather err on the side of caution and encourage their continued use, stressing that there is no room for complacency when it comes to ensuring patient safety.[25] This opinion is similarly echoed by the National Institute for Health and Care Excellence guidelines which assert that mask usage contributes towards 'maintaining theatre discipline'.

Another unavoidable aspect of this debate is that of public perception. In the public psyche, facemasks have become so strongly associated with safe and proper surgical practice that their disposal could cause unnecessary patient distress. Indeed, the response on various medical forums following Mr Ahmed's decision not to wear a mask during his broadcasted surgeries would reflect the prevalence of such a belief among the public.

It is clear that more studies are required before any absolute conclusions can be drawn regarding the effectiveness or, indeed, ineffectiveness of surgical masks. The published literature does suggest that it may be reasonable to further examine the need for masks in contemporary surgical practice given the interests of comfort, budget constraints and potential ease of communication, although any such study would undoubtedly have to be large and well controlled to prove causality given the low event frequency of surgical site infections. It is possible, if not probable, that if surgical facemasks were to be introduced today, without the historical impetus currently associated with their use, the experimental evidence would not be sufficiently compelling to incorporate facemasks into surgical practice.

However, when current surgical practice is the culmination of layer upon layer of precautions in the hope of preventing surgical site infection, do we dare to experiment with their omission to see if they have any tangible consequence on morbidity and mortality? A randomised control trial investigating the uncertainty surrounding prophylactic antibiotic use in clean coronary artery surgery turned out to be catastrophic – the study had to be terminated early for ethical reasons due to an unacceptable increase in postoperative infection in the placebo cohort.[26] Perhaps an annual expenditure of a few million pounds in a healthcare budget of almost £100 billion is a small price to pay for an intervention of unknown but potentially dramatic effectiveness.

It is important not to construe an absence of evidence for effectiveness with evidence for the absence of effectiveness. While there is a lack of evidence supporting the effectiveness of facemasks, there is similarly a lack of evidence supporting their ineffectiveness. With the information currently available, it would be imprudent to recommend the removal of facemasks from surgery. Instead, in the medical field where common practice can so easily become dogma, it is necessary to recognise the constant need to maintain a healthy scepticism towards established beliefs and to periodically re-evaluate and critically assess their scientific merit.

Declarations

Competing interests: None declared.

Funding: None declared.

Ethical approval: Not applicable.

Guarantor: CDZ.

Contributorship: CDZ drafted the paper and all other authors commented on the paper and contributed to amendments.

Acknowledgements: None.

Provenance: Not commissioned; peer-reviewed by Tom Treasure.

References

1. Leyland M and McCloy R. Surgical face masks: protection of self or patient? *Ann R Coll Surg Engl* 1993; 75: 1.

2. Spooner JL. History of surgical face masks. *AORN J* 1967; 5: 76–80. See http://www.aornjournal.org/article/S0001-2092(08)71359-0/abstract (last checked 31 March 2015).
3. Mikulicz J. Das Operieren in Sterilisirten Zwirnhandschuhen und Mit Mundbinde. *Cent f Chir* 1897; 25: 714–717.
4. Järvinen TLN, Sievänen H, Kannus P, Jokihaara J and Khan KM. The true cost of pharmacological disease prevention. *BMJ* 2011; 342: d2175.
5. Belkin NL. A century after their introduction, are surgical masks necessary? *AORN J* 1996; 64: 602. See http://www.aornjournal.org/article/S0001-2092(06)63628-4/abstract (last checked 31 March 2015).
6. Schweizer RT. Mask wiggling as a potential cause of wound contamination. *Lancet* 1976; 2: 1129–1130.
7. Orr NW. Is a mask necessary in the operating theatre? *Ann R Coll Surg Engl* 1981; 63: 390–392.
8. National Collaborating Centre for Women's and Children's Health. *Surgical Site Infection: Prevention and Treatment of Surgical Site Infection*. UK: Nursing standard (Royal College of Nursing 1987), 2008. See www.rcog.org.uk (last checked 31 March 2015).
9. Lipp A and Edwards P. Disposable surgical face masks for preventing surgical wound infection in clean surgery. *Cochrane Database Syst Rev* 2002; 1: CD002929.
10. Chamberlain GV and Houang E. Trial of the use of masks in the gynaecological operating theatre. *Ann R Coll Surg Engl* 1984; 66: : 432–433.
11. Tunevall TG. Postoperative wound infections and surgical face masks: a controlled study. *World J Surg* 1991; 15: 383–387. DOI: 10.1007/BF01658736.
12. Lipp A and Edwards P. Disposable surgical face masks for preventing surgical wound infection in clean surgery. *Cochrane Database Syst Rev* 2014; 2: CD002929.
13. Webster J, Croger S, Lister C, Doidge M, Terry MJ and Jones I. Use of face masks by non-scrubbed operating room staff: a randomized controlled trial. *ANZ J Surg* 2010; 80: 169–173.
14. Filochowski J. FOI Reply 560, 2012. p. 3. See https://www.westhertshospitals.nhs.uk/foi_publication_scheme/disclosure_log/2012/may/documents/FOI Reply 560.pdf (last checked 30 October 2014).
15. Health and Social Care Information Centre. *Hospital Episode Statistics Database*. See http://www.hscic.gov.uk/hesdata (last checked 30 October 2014).
16. NHS Procurement Atlas of Variation. See http://ccgtools.england.nhs.uk/procurement/ProcAtlasJuly2014/atlas.html (last checked 30 October 2014).
17. Plowman R, Graves N, Griffin MAS, Roberts JA, Swan AV, Cookson B, et al. The rate and cost of hospital-acquired infections occurring in patients admitted to selected specialties of a district general hospital in England and the national burden imposed. *J Hosp Infect* 2001; 47: 198–209.
18. Siegel JD, Rhinehart E, Jackson M and Chiarello L. 2007 guideline for isolation precautions: preventing transmission of infectious agents in health care settings. *Am J Infect Control* 2007; 35: S65–S164.
19. Davies CG, Khan MN, Ghauri ASK and Ranaboldo CJ. Blood and body fluid splashes during surgery – the need for eye protection and masks. *Ann R Coll Surg Engl* 2007; 89: 770–772.
20. Aisien AO and Ujah IAO. Risk of blood splashes to masks and goggles during cesarean section. *Med Sci Monit* 2006; 12: CR94–CR97.
21. Weber A, Willeke K, Marchioni R, Myojo T, McKay R, Donnelly J, et al. Aerosol penetration and leakage characteristics of masks used in the health care industry. *Am J Infect Control* 1993; 21: 167–173.
22. Ransjö U. Masks: a ward investigation and review of the literature. *J Hosp Infect* 1986; 289–294.
23. Reingold AL, Kane MA and Hightower AW. Failure of gloves and other protective devices to prevent transmission of hepatitis B virus to oral surgeons. *JAMA* 1988; 259: 2558–2560.
24. Glasziou P, Chalmers I, Rawlins M and McCulloch P. When are randomised trials unnecessary? Picking signal from noise. *BMJ* 2007; 334: 349–351.
25. Hogan B and Samaranayake LP. The surgical mask unmasked: a review. *Oral Surgery, Oral Med Oral Pathol* 1990; 70: :–34–36.
26. Penketh ARL, Wansbrough-Jones MH, Wright E, Imrie F, Pepper JR and Parker DJ. Antibioitic prophylaxis for coronary artery bypass graft surgery. *Lancet* 1985; 325: 1500. See http://www.sciencedirect.com/science/article/pii/S0140673685922676 (last checked 31 March 2015).

Appendix 1. Databases included in the search strategy[12]

The Cochrane Wounds Group Register (searched 23 October 2013)
The Cochrane Central Register of Controlled Trials (CENTRAL) (*The Cochrane Library 2013*, Issue 9)
Ovid MEDLINE (1946 to October Week 3 2013)
Ovid MEDLINE (In-process and other non-indexed citations, October 23, 2013)
Ovid EMBASE (1974 to 23 October 2013)
EBSCO CINAHL (1982 to 18 October 2013)

Randomized Controlled Trial PLoS One. 2010 Nov 17;5(11):e13998.
doi: 10.1371/journal.pone.0013998.

Surgical mask to prevent influenza transmission in households: a cluster randomized trial

Laetitia Canini [1], Laurent Andréoletti, Pascal Ferrari, Romina D'Angelo, Thierry Blanchon, Magali Lemaitre, Laurent Filleul, Jean-Pierre Ferry, Michel Desmaizieres, Serge Smadja, Alain-Jacques Valleron, Fabrice Carrat

Affiliations

PMID: 21103330 PMCID: PMC2984432 DOI: 10.1371/journal.pone.0013998
Free PMC article

Abstract

Background: Facemasks and respirators have been stockpiled during pandemic preparedness. However, data on their effectiveness for limiting transmission are scarce. We evaluated the effectiveness of facemask use by index cases for limiting influenza transmission by large droplets produced during coughing in households.

Methodology and principal findings: A cluster randomized intervention trial was conducted in France during the 2008-2009 influenza season. Households were recruited during a medical visit of a household member with a positive rapid influenza A test and symptoms lasting less than 48 hours. Households were randomized either to the mask or control group for 7 days. In the intervention arm, the index case had to wear a surgical mask from the medical visit and for a period of 5 days. The trial was initially intended to include 372 households but was prematurely interrupted after the inclusion of 105 households (306 contacts) following the advice of an independent steering committee. We used generalized estimating equations to test the association between the intervention and the proportion of household contacts who developed an influenza-like illness during the 7 days following the inclusion. Influenza-like illness was reported in 24/148 (16.2%) of the contacts in the intervention arm and in 25/158 (15.8%) of the contacts in the control arm and the difference between arms was 0.40% (95%CI: -10% to 11%, P = 1.00). We observed a good adherence to the intervention. In various sensitivity analyses, we did not identify any trend in the results suggesting effectiveness of facemasks.

Conclusion: This study should be interpreted with caution since the lack of statistical power prevents us to draw formal conclusion regarding effectiveness of facemasks in the context of a seasonal epidemic.

Trial registration: clinicaltrials.gov NCT00774774.

Figures

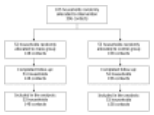


Figure 1. Flow diagram.

Related information

Cited in Books
MedGen

LinkOut - more resources

Full Text Sources

Europe PubMed Central
PubMed Central
Public Library of Science

Medical

ClinicalTrials.gov
MedlinePlus Health Information

Randomized Controlled Trial Am J Infect Control. 2009 Jun;37(5):417-419.
doi: 10.1016/j.ajic.2008.11.002. Epub 2009 Feb 12.

Use of surgical face masks to reduce the incidence of the common cold among health care workers in Japan: a randomized controlled trial

Joshua L Jacobs [1], Sachiko Ohde [2], Osamu Takahashi [3], Yasuharu Tokuda [3], Fumio Omata [3], Tsuguya Fukui [3]

Affiliations
PMID: 19216002 DOI: 10.1016/j.ajic.2008.11.002

Abstract

Background: Health care workers outside surgical suites in Asia use surgical-type face masks commonly. Prevention of upper respiratory infection is one reason given, although evidence of effectiveness is lacking.

Methods: Health care workers in a tertiary care hospital in Japan were randomized into 2 groups: 1 that wore face masks and 1 that did not. They provided information about demographics, health habits, and quality of life. Participants recorded symptoms daily for 77 consecutive days, starting in January 2008. Presence of a cold was determined based on a previously validated measure of self-reported symptoms. The number of colds between groups was compared, as were risk factors for experiencing cold symptoms.

Results: Thirty-two health care workers completed the study, resulting in 2464 subject days. There were 2 colds during this time period, 1 in each group. Of the 8 symptoms recorded daily, subjects in the mask group were significantly more likely to experience headache during the study period ($P < .05$). Subjects living with children were more likely to have high cold severity scores over the course of the study.

Conclusion: Face mask use in health care workers has not been demonstrated to provide benefit in terms of cold symptoms or getting colds. A larger study is needed to definitively establish noninferiority of no mask use.

Related information

Cited in Books
MedGen

LinkOut – more resources

Full Text Sources
ClinicalKey

ClinicalKey Nursing
Elsevier Science
Ovid Technologies, Inc.

Medical

MedlinePlus Health Information

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

International Journal of Nursing Studies 108 (2020) 103629

Contents lists available at ScienceDirect

International Journal of Nursing Studies

journal homepage: www.elsevier.com/ijns

A rapid systematic review of the efficacy of face masks and respirators against coronaviruses and other respiratory transmissible viruses for the community, healthcare workers and sick patients

C. Raina MacIntyre [a], Abrar Ahmad Chughtai [b,*]

[a] The Kirby Institute, UNSW Medicine, University of New South Wales, Sydney, Australia
[b] School of Public Health and Community Medicine, UNSW Medicine, University of New South Wales, Room 228, Level 2 Samuels Building, Sydney, Australia

ARTICLE INFO

Article history:
Received 24 March 2020
Received in revised form 18 April 2020
Accepted 21 April 2020

Keywords:
Coronavirus
Coronavirus disease
COVID19
Mask
Respirators
Personal protective equipment

ABSTRACT

Background: The pandemic of COVID-19 is growing, and a shortage of masks and respirators has been reported globally. Policies of health organizations for healthcare workers are inconsistent, with a change in policy in the US for universal face mask use. The aim of this study was to review the evidence around the efficacy of masks and respirators for healthcare workers, sick patients and the general public.

Methods: A systematic review of randomized controlled clinical trials on use of respiratory protection by healthcare workers, sick patients and community members was conducted. Articles were searched on Medline and Embase using key search terms.

Results: A total of 19 randomised controlled trials were included in this study – 8 in community settings, 6 in healthcare settings and 5 as source control. Most of these randomised controlled trials used different interventions and outcome measures. In the community, masks appeared to be effective with and without hand hygiene, and both together are more protective. Randomised controlled trials in health care workers showed that respirators, if worn continually during a shift, were effective but not if worn intermittently. Medical masks were not effective, and cloth masks even less effective. When used by sick patients randomised controlled trials suggested protection of well contacts.

Conclusion: The study suggests that community mask use by well people could be beneficial, particularly for COVID-19, where transmission may be pre-symptomatic. The studies of masks as source control also suggest a benefit, and may be important during the COVID-19 pandemic in universal community face mask use as well as in health care settings. Trials in healthcare workers support the use of respirators continuously during a shift. This may prevent health worker infections and deaths from COVID-19, as aerosolisation in the hospital setting has been documented.

What is already known about the topic?

- Masks and respirators are commonly used to protect from respiratory infections in three different indications – for healthcare workers, sick patients and well community members.
- Currently there is debate and conflicting guidelines around the use of masks and respirators in healthcare and community settings.

What this paper adds

- In the community, masks may be more protective for well people.
- In healthcare settings continuous use of respirators, is more protective compared to the medical masks, and medical masks are more protective than cloth masks. Depending on the fabric and design, some cloth masks may not be safe for healthcare workers.
- The use of masks by sick patients is likely protective, and coronaviruses can be emitted in normal breathing, in fine airborne particles.

* Corresponding author.
E-mail addresses: r.macintyre@unsw.edu.au (C.R. MacIntyre), abrar.chughtai@unsw.edu.au (A.A. Chughtai).

https://doi.org/10.1016/j.ijnurstu.2020.103629

1. Introduction

The use of personal protective equipment for coronavirus disease (COVID-19) has been controversial, with differing guidelines issued by different agencies (Chen et al., 2020). COVID-19 is caused by severe acute respiratory syndrome coronavirus2 (SARS-CoV-2), a beta-coronavirus, similar to severe acute respiratory syndrome coronavirus (SARS CoV) (Chen et al., 2020). Seasonal alpha and beta coronaviruses cause common colds, croup and bronchiolitis. The transmission mode of coronaviruses in humans is similar, thought to be by droplet, contact and sometimes airborne routes (Ong et al., 2020; Zhang et al., 2020; Zou et al., 2020). The World Health Organization recommends surgical mask for health workers providing routine care to a coronavirus disease patient (World Health Organisation (WHO) 2020), whilst the US Centers for Disease Control and Prevention recommended a respirator (Center for Disease Control and Prevention CDC, 2020). Most authorities, except the US CDC, are recommending that community members not wear a mask, and that a mask should only be worn by a sick patient (also referred to as *source control*) (Chughtai et al., 2020). There are more randomised controlled trials of community use of masks in well people than studies of the use by sick people (*source control*). The aim of this study was to review the randomised controlled trials evidence for use of masks and respirators by the community, health care workers and sick patients for prevention of infection.

2. Methods

We searched Medline and EmBase for clinical trials on masks and respirators using the key words “mask”, “respirator”, and “personal protective equipment”. The search was conducted between 1 March to April 17 2020, and all randomised controlled trials published before the search date were included. Two authors (CRM and AAC) reviewed the title and abstracts to identify randomised controlled trials on masks and respirators. We also searched relevant papers from the reference lists of previous clinical trials and systematic reviews. Studies that were not randomised controlled trials, were about anesthesia, or not about prevention of infection were excluded. Animal studies, experimental and observational epidemiologic studies were also excluded. Studies published in English language were included.

We found 602 papers on Medline and 250 on Embase. 820 papers were excluded by title and abstract review. Full texts were reviewed for 32 papers and 19 were selected in this review. Results were reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria (Moher et al., 2015).

3. Results

In general, the results show protection for healthcare workers and community members, and likely benefit of masks used as source control. We found eight clinical trials (Aiello et al., 2012; Simmerman et al., 2011; Larson et al., 2010; Aiello et al., 2010; MacIntyre et al., 2009; Cowling et al., 2008, Suess et al., 2012; Cowling et al., 2009) on the use of masks in the community (Table 1). In the community, masks appear to be effective with and without hand hygiene, and both together are more protective (Aiello et al., 2012; Aiello et al., 2010; MacIntyre et al., 2009). However, some randomised controlled trials which measured both hand hygiene and masks measured the effect of hand hygiene alone, but not of masks alone (Simmerman et al., 2011, Cowling et al., 2009). In more than one trial, interventions had to be used within 36 hours of exposure to be effective (Cowling et al., 2009; Suess et al., 2012).

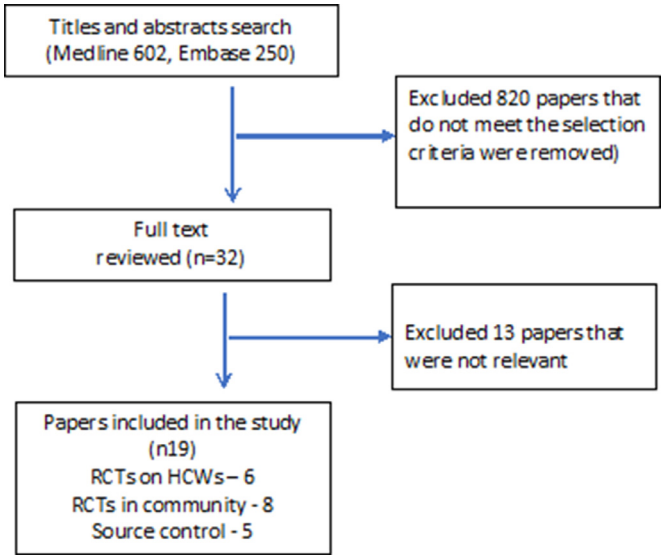


Fig. 1. Search strategy and selection of papers.

To date, six randomised controlled trials (Radonovich et al., 2019; Jacobs et al., 2009, Loeb et al., 2009; MacIntyre et al., 2011, 2013, 2015) have been conducted on the use of masks and/or respirators by healthcare workers in health care settings (Table 2). The healthcare worker trials (Table 2) used different interventions and different outcome measures, and one was in the outpatient setting. A Japanese study had only 32 subjects, and likely was underpowered to find any difference between masks and control (Jacobs et al., 2009). Two North American trials of masks and respirators against influenza infection found no difference between the arms, but neither had a control arm to differentiate equal efficacy from equal inefficacy (Radonovich et al., 2019, Loeb et al., 2009). Neither trial can prove equivalence, as this requires one intervention to be already proven efficaceous against placebo. Without a control group to determine rates of influenza in unprotected healthcare workers, neither study is able to determine efficacy if no difference was observed between the two interventions. A serologic study showed that up to 23% of unprotected healthcare workers (a rate identical to that observed in Loeb the trial, which also used serology) contract influenza during outbreaks (Elder et al., 1996), which suggests lack of efficacy. Studies of nosocomial influenza generally find lower influenza attack rates in unprotected healthcare workers than observed in the Loeb trial (Salgado et al., 2002).

Further problems with this study are that the majority of subjects were defined as having influenza on the basis of serological positivity (Loeb et al., 2009). The 10% seroconversion to pandemic H1N109 (with no pandemic virus isolation or positive PCR) observed in the trial, suggests that pandemic H1N109 was circulating in Ontario before April 2009, which is unlikely.

A serological definition of influenza can be affected by vaccination. The authors claim they excluded influenza vaccinated subjects in the outcome, but according to figure 1 in the Loeb trial, (Loeb et al., 2009) these subjects (130 in total) are included in the analysis. If they had been excluded and even if no other subjects were excluded, the total analysed would be 348, which is lower than the 422 subjects analysed (Loeb et al., 2009). These 130 vaccinated subjects should have been excluded entirely from the analysis. The vaccination status of subjects with seropositivity is not provided in the paper, but it appears people with positive serology due to vaccination may have been misclassified as influenza cases (Loeb et al., 2009).

Table 1
Community mask trials.

Author, year	N, country	Interventions	Results
Cowling et al. (2008)	198 Households Hong Kong	Medical masks Hand washing Control	NS – this was a preliminary report of the 2009 trial.
MacIntyre et al. (2009)	143 Households Australia	Medical masks P2 masks Control	Intention to treat non-significant. Adherence with mask wearing low (25-30% by day 5). In sub-analysis, masks/P2 protective if adherent.
Cowling et al., 2009)	407 households Hong Kong	Hand hygiene Masks + hand hygiene Control	Intention to treat not significant. Masks plus hand hygiene protective against lab confirmed influenza if used within 36 hours. Hand hygiene alone not significant.
Aiello et al. (2010)	1437 college students, United States of America	Masks Masks + hand washing Control	Intention to treat non-significant. Masks + handwashing protective in week 4 -6 of observation and beyond.
Aiello et al. (2012)	1178 college students, United States of America	Masks Masks + hand hygiene Control	Intention to treat non-significant. Masks + hand hygiene protective in week 3 of observation and beyond. Masks alone not protective.
Larson et al. (2010)	617 households, United States of America	Health education (HE) Hand hygiene + HE Masks + hand hygiene +HE	Masks + hand hygiene + HE protective against secondary transmission measured by confirmed influenza and ILI. Mean secondary attack rates for HE, HE + HH, HE+HH+M groups were 0.023, 0.020, and 0.018, respectively
Simmerman et al. (2011)	465 index patients and their families, Thailand	Hand hygiene Masks + hand hygiene Control	No significant difference in confirmed influenza infection
Suess et al. (2012)	84 index cases and 218 household contacts, Germany	Masks Masks + hand hygiene Control	Intention to treat analysis was non-significant. Where used within 36 h, secondary infection in the pooled M and MH groups was significantly lower compared to the control group. In multivariable analysis for predictors of qRT-PCR confirmed influenza infection and clinical influenza among included households in separate models allowing for within household correlation, M and MH were protective against Influenza AH1N1pdm09.

Table 2
Trials of mask and respirator use by health care workers.

Author, year	N healthcare workers, Country	Interventions	Results
Jacobs et al. (2009)	32 Japan	Medical masks Control	NS
Loeb et al. (2009)	446 Canada	Medical masks, targeted N95	No significant difference between Masks and targeted N95
MacIntyre et al. (2011)	1441 China	Masks N95 respirators, fit tested N95 respirators, non-fit tested Control	Continuous N95 protective against clinical, viral and bacterial endpoints
MacIntyre et al. (2013)	1669 China	Medical Mask N95 (continuous) N95 (targeted)	Continuous N95 protective No difference between targeted N95 and medical masks
MacIntyre et al. (2015)	1607 Vietnam	Medical masks, cloth masks, control	Medical masks protective or Cloth masks increase risk of infection
Radonovich et al. (2019)	2862 United States of America	Medical masks, targeted N95 (when 2 m from confirmed respiratory infection) in Outpatient setting.	No significant difference between Masks and targeted N95

In both the North American trials, the intervention comprised wearing the mask or respirator when in contact with recognized ILI or when doing a high risk procedure, which is a targeted strategy (Radonovich et al., 2019, Loeb et al., 2009). One was in an outpatient setting. (Radonovich et al., 2019) We conducted a randomised controlled trial comparing the targeted strategy tested in the two North American studies, with the wearing of respiratory protection during an entire shift, and showed efficacy for continual (but not targeted) use of a respirator (MacIntyre et al., 2013). The study also did not show efficacy for a surgical mask worn continually, and therefore no difference between a surgical mask and targeted use of a respirator (MacIntyre et al., 2013), which is consistent with the findings of the North American trials (Radonovich et al., 2019, Loeb et al., 2009). In summary, the evidence is consistent that a respirator must be worn throughout the shift to be protective. Targeted use of respirators only when doing high risk procedures and medical mask use is not protective. Another randomised controlled trial we conducted in China showed efficacy for continual use of a respirator, but not for a mask, and also found fit-testing of the respirator did not affect efficacy (MacIntyre et al., 2011). However, this may be specific to the quality of the tested product, and is not generalisable to other respirators – fit testing is a necessary part of respirator use (Chughtai et al., 2015).

For healthcare workers, there is evidence of efficacy of respirators if worn continually during a shift, but no evidence of efficacy of a mask (MacIntyre et al., 2011, 2013). For hospitals where COVID-19 patients are being treated, there is growing evidence of widespread contamination of the ward environment, well beyond 2 m from the patient, as well as aerosol transmission (Ong et al., 2020; Santarpia et al., n.d.; Guo et al., 2020). Several studies have

Table 3
Trials of Masks used by a sick patient as source control.

Author, year	*N*, country	Interventions	Results
Johnson et al. (2009)	9 subjects with confirmed influenza, Australia	Medical mask N95 (participants coughed 5 times onto a Petri dish wearing each device)	NS - Surgical and N95 masks were equally effective in preventing the spread of PCR-detectable influenza
Canini et al. (2010)	105 index cases and 306 household contacts, France	Medical mask Control	No significant difference, but trial terminated early
MacIntyre et al. (2016)	245 index cases and 597 household contacts,	Medical mask worn by sick case Control (no mask) Household contacts Followed for infection.	Intention to treat analysis not significant. Mask protective if worn
Barasheed et al. (2014)	Hajj Setting. 22 tents were randomised to 'mask' ($n = 12$) or 'control' ($n = 10$) 75 pilgrims in 'mask' and 89 in 'control' group Saudi Arabia	Mask and control	Less ILI among the contacts of mask users compared to the control tents (31% versus 53%, $p = 0.04$). Laboratory results did not show any difference between the two groups
Leung et al. (2020)	Experimental study of 246 subjects randomised to surgical mask and no mask	Mask and control	111 were infected by human (seasonal) coronavirus. Coronavirus found in exhaled breath of no-mask subjects but not in mask wearers. More virus was found in fine aerosols than large droplets

found SARS-CoV-2 on air vents and in air samples in intensive care units and COVID-19 wards (Santarpia et al., n.d.; Chia et al., 2020; Liu et al., 2020), and an experimental study showed the virus in air samples three hours after aerosolization (van Doremalen et al., 2020). The weight of this evidence and the precautionary principle (MacIntyre et al., 2014a; 2014b), favors respirators for healthcare workers. We showed lower rates of infection outcomes in the medical mask arm compared to control, but the difference was not significant (MacIntyre et al., 2011). It could be that larger trials are needed to demonstrate efficacy of a mask, but any protection is far less than from a respirator. A trial we conducted in Vietnam of 2-layered cotton cloth masks compared to medical masks showed a lower rate of infection in the medical mask group, and a 13 times higher risk of infection in the cloth mask arm (MacIntyre et al., 2015). The study suggests cloth masks may increase the risk of infection (MacIntyre et al., 2015), but may not be generalizable to all homemade masks. The material, design and adequacy of washing of cloth masks may have been a factor (Macintyre et al., 2020). There are no other randomised controlled trial of cloth masks published at this time, but if any protection is offered by these it would be less than even a medical mask.

Table 3 shows the trials of source control. There were five randomised controlled trials identified of masks used by sick patients (Johnson et al., 2009, Barasheed et al., 2014; Leung et al., 2020; MacIntyre et al., 2016; Canini et al., 2010). One was an experimental study of 9 influenza patients, which did not measure clinical endpoints (Johnson et al., 2009). Participants with confirmed influenza coughed onto culture medium wearing a N95 respirator or a mask. No influenza grew on the medium. A trial of 105 sick patients wearing a mask (or no mask) in the household found no significant difference between arms (Canini et al., 2010). However, the trial was terminated prematurely and did not meet recruitment targets, so was probably underpowered. One randomised controlled trial was conducted among Hajj pilgrims, with both well and sick pilgrims wearing masks, and low rates of ILI were reported among contact of mask pilgrims (Barasheed et al., 2014). Our randomised controlled trial is the largest available with clinical endpoints, and studied 245 patients randomised to mask or control (MacIntyre et al., 2016). Compliance was suboptimal in the mask group and some controls wore masks. The intention to treat analysis showed no difference, but when analysed by actual mask use, the rate of infection in household contacts was lower in those who wore masks (MacIntyre et al., 2016). A trial with an experimental design was published in April 2020, examining a range of viruses including seasonal human coronaviruses (Leung et al., 2020). This showed that coronaviruses are preferentially found in aerosolized particles compared to large droplets, and could be expelled by normal tidal breathing. Wearing a surgical mask prevented virus from being exhaled.

4. Discussion

There are more randomised controlled trials of community use of masks in well people (Aiello et al., 2012; Simmerman et al., 2011; Larson et al., 2010; Aiello et al., 2010; MacIntyre et al., 2009; Cowling et al., 2008, Suess et al., 2012, Cowling et al., 2009) than studies of the use by sick people (also referred to as "*source control*"), and these trials are larger than the few on source control (Johnson et al., 2009, Leung et al., 2020; MacIntyre et al., 2016). The evidence suggests protection by masks in high transmission settings such as household and college settings, especially if used early, in some trials if combined with hand hygiene and if wearers are compliant (Aiello et al., 2012; Aiello et al., 2010; MacIntyre et al., 2009; Cowling et al., 2008, 2009; Suess et al., 2012). If masks protect in high transmission settings, they should also protect in crowded public spaces, including workplaces, buses, trains, planes and other closed settings. The trial which did not show efficacy used influenza as the outcome measure (Simmerman et al., 2011), which is a rare outcome, so requires a larger sample size for adequate power and may have been underpowered.

For healthcare workers, the only trials to show a difference between respirators and masks demonstrated efficacy for continuous use of a respirator through a clinical shift, but not masks (MacIntyre et al., 2011, 2013). The two trials which showed no difference are widely cited as evidence that masks provide equal protection as respirators (Radonovich et al., 2019, Loeb et al., 2009). However, without a control arm, the absence of difference between arms could reflect equal efficacy or inefficacy, and it is not possible to draw any conclusions about efficacy. The outpatient setting in the US trial may have had lower exposure risk than the inpatient setting of other trials. (Radonovich et al., 2019) In both the North American trials, the intervention comprised wearing the mask or respirator intermittently when in contact with recognized ILI or when doing a high risk procedure (Radonovich et al., 2019, Loeb et al., 2009). The underlying assumption that the majority of infections in healthcare workers occur during self-identified high-risk exposures is not supported by any evidence. It assumes healthcare workers can accurately identify when they are risk in a busy, clinical setting, when the majority of infections may occur when healthcare workers are unaware of the risk (such as when walking

through a busy emergency room or ward where aerosolized virus may be present). Conversely, infections could occur outside the workplace. This could explain the lack of difference if there was no actual efficacy of either arm and if much of the infection occurs in unrecognized situations of risk either within or outside the workplace.

In practice, hospital infection control divides infections into droplet or airborne spread, and recommends droplet (mask) or airborne (respirator) precautions accordingly (MacIntyre et al., 2017). In a pooled analysis of both healthcare worker trials, we showed that continual use of a respirator is more efficacious in protecting healthcare workers even against infections assumed to be spread by the droplet route (MacIntyre et al., 2017). Medical masks did not significantly protect against viral, bacterial, droplet or other infection outcomes. However, the summary odds ratio for masks was less than one, which suggests a low level of protection. Targeted use of respirator protected against bacterial and droplet infections, but not against viral infections, suggesting viral infections may be more likely to be airborne in the hospital setting (MacIntyre et al., 2017).

The five available studies of mask use by sick patients suggest a benefit, but are much smaller trials than the community trials, two without clinical endpoints, and with less certainty around the findings (Johnson et al., 2009, Barasheed et al., 2014; Leung et al., 2020; MacIntyre et al., 2016; Canini et al., 2010). Only 3/5 trials examined clinical outcomes in close contacts (Barasheed et al., 2014; MacIntyre et al., 2016; Canini et al., 2010) and suggest a benefit

Many systematic reviews have been conducted on masks, respirators and other PPE in past (Cowling et al., 2010; Bin-Reza et al., 2012; Gralton and McLaws, 2010; Gamage et al., 2005; Jefferson et al., 2009; Jefferson et al., 2011; Jefferson et al., 2008; Aledort et al., 2007; Lee et al., 2011; Verbeek et al., 2020). These reviews generally examined multiple interventions (e.g. masks and hand hygiene etc.), often combined different outcome measures that were not directly comparable and were inconclusive. Moreover, most of these reviews did not include more recent randomised controlled trials (Radonovich et al., 2019, MacIntyre et al., 2015). This systematic review only focuses on masks and respirators and contains all new studies.

In summary, there is a growing body of evidence supporting all three indications for respiratory protection – community, healthcare workers and sick patients (source control). The largest number of randomised controlled trials have been done for community use of masks by well people in high-transmission settings such as household or college settings. There is benefit in the community if used early, with hand hygiene and if compliant.

Respirators protect healthcare workers if worn continually, but not if worn intermittently in self-identified situations of risk. This supports the suggestion that the health care environment is a risk to healthcare workers even when not doing aerosol generating procedures or caring for a known infectious patient. For COVID-19 specifically, the growing body of evidence showing aerosolisation of the virus in the hospital ward highlights the risk of inadvertent exposure for healthcare workers and supports the use of airborne precautions at all times on the ward (Santarpia et al., n.d.; Chia et al., 2020; Liu et al., 2020). Further, the rule of 1–2 m of spatial separation is not based on good evidence, with most research showing that droplets can travel further than 2 m, and that infections cannot be neatly separated into droplet and airborne (MacIntyre et al., 2017; Bahl et al., 2020). In the UK, one healthcare trust found almost one in five healthcare workers to be infected with COVID-19 (Keeley et al., 2020). The deaths of healthcare workers from COVID-19 reflect this risk (Zhan et al., 2020). The use of masks by sick people, despite being the WHO's only recommendation for mask use by community members during COVID-19 pandemic, is supported by the smallest body of evidence. Source control is probably a sensible recommendation given the suggestion of protection and given specific data on coronaviruses showing protection (Leung et al., 2020). It may help if visitors and febrile patients wear a mask in the healthcare setting, whether in primary care or hospitals. Universal face mask use is likely to have the most impact on epidemic growth in the community, given the high risk of asymptomatic and pre-symptomatic transmission (He et al., 2020).

Conflict of Interest

C Raina MacIntyre receives funding from NHMRC (centre for Research Excellence and Principal Research Fellowship) and Sanofi currently. She has received funding from 3M more than 10 years ago for face mask research.

Abrar Ahmad Chughtai had testing of filtration of masks by 3M for his Ph.D. more than 10 years ago. 3M products were not used in his research. He also has worked with CleanSpace Technology on research on fit testing of respirators (no funding was involved).

Funding

C Raina MacIntyre is supported by a NHMRC Principal Research Fellowship, grant number 1137582.

References

Aiello, A.E., Perez, V., Coulborn, R.M., Davis, B.M., Uddin, M., Monto, A.S., 2012. Facemasks, hand hygiene, and influenza among young adults: a randomized intervention trial. PLoS ONE 7 (1).

Aiello, A.E., Murray, G.F., Perez, V., Coulborn, R.M., Davis, B.M., Uddin, M., et al., 2010. Mask use, hand hygiene, and seasonal influenza-like illness among young adults: a randomized intervention trial. J. Infect. Dis. 201 (4), 491–498.

Aledort, J.E., Lurie, N., Wasserman, J., Bozzette, S.A., 2007. Non-pharmaceutical public health interventions for pandemic influenza: an evaluation of the evidence base. BMC Public Health 7, 208.

Bahl, P., Doolan, C., de Silva, C., Chughtai, A.A., Bourouiba, L., MacIntyre, C.R., 2020. Airborne or droplet precautions for health workers treating COVID-19? J. Infect. Dis..

Barasheed, O., Almasri, N., Badahdah, A.M., Heron, L., Taylor, J., McPhee, K., et al., 2014. Pilot randomised controlled trial to test effectiveness of facemasks in preventing influenza-like illness transmission among Australian Hajj Pilgrims in 2011. Infect. Disord. Drug Targets. 14 (2), 110–116.

Bin-Reza, F., Lopez Chavarrias, V., Nicoll, A., Chamberland, M.E., 2012. The use of masks and respirators to prevent transmission of influenza: a systematic review of the scientific evidence. Influenza Other Respir. Viruses 6 (4), 257–267.

Canini L., Andreoletti L., Ferrari P., D'Angelo R., Blanchon T., Lemaitre M., et al. Surgical mask to prevent influenza transmission in households: a cluster randomized trial. PLoS ONE [Electronic Resource]. 2010. 5(11):e13998.

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., et al., 2020. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. The Lancet 395 (10223), 507–513.

Center for Disease Control and Prevention (CDC). Interim healthcare infection prevention and control recommendations for patients under investigation for 2019 novel coronavirus. 2020. January 2020 [Available from: https://www.cdc.gov/coronavirus/2019-nCoV/infection-control.html.

Chughtai, A.A., Seale, H., Islam, M.S., Owais, M., Macintyre, C.R., 2020. Policies on the use of respiratory protection for hospital health workers to protect from coronavirus disease (COVID-19). Int. J. Nurs. Stud. 105, 103567.

Chia, P.Y., Coleman, K.K., Tan, Y.K., Ong, S.W.X., Gum, M., Lau, S.K., et al., 2020. Detection of Air and Surface Contamination by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in Hospital Rooms of Infected Patients. medRxiv doi:10.1101/2020.03.29.20046557.

Chughtai, A.A., MacIntyre, C.R., Zheng, Y., Wang, Q., Toor, Z.I., Dung, T.C., et al., 2015. Examining the policies and guidelines around the use of masks and respirators by healthcare workers in China, Pakistan and Vietnam. J. Infect. Prev. 16 (2), 68–74.

Cowling, B.J., Chan, K.H., Fang, V.J., Cheng, C.K., Fung, R.O., Wai, W., et al., 2009. Facemasks and hand hygiene to prevent influenza transmission in households: a cluster randomized trial. Ann. Intern. Med. 151 (7), 437–446.

Cowling, B.J., Fung, R.O., Cheng, C.K., Fang, V.J., Chan, K.H., Seto, W.H., et al., 2008. Preliminary findings of a randomized trial of non-pharmaceutical interventions to prevent influenza transmission in households. PLoS ONE 3 (5), e2101.

Cowling, B.J., Zhou, Y., Ip, D.K., Leung, G.M., Aiello, A.E., 2010. Face masks to prevent transmission of influenza virus: a systematic review. Epidemiol. Infect. 138 (4), 449–456.

Elder, A.G., O'Donnell, B., McCruden, E.A., Symington, I.S., Carman, W.F., 1996. Incidence and recall of influenza in a cohort of Glasgow healthcare workers during the 1993-4 epidemic: results of serum testing and questionnaire. BMJ (Clin. Res. Ed.) 313 (7067), 1241–1242.

Gamage, B., Moore, D., Copes, R., Yassi, A., Bryce, E., 2005. Protecting health care workers from SARS and other respiratory pathogens: a review of the infection control literature. Am. J. Infect. Control 33 (2), 114–121.

Gralton, J., McLaws, M.L., 2010. Protecting healthcare workers from pandemic influenza: N95 or surgical masks? Crit. Care Med. 38 (2), 657–667.

Guo, Z., Wang, Z., Zhang, S., Li, X., Li, L., Li, C., et al., 2020. Aerosol and Surface Distribution of Severe Acute Respiratory Syndrome Coronavirus 2 in Hospital Wards, Wuhan, China, 2020. Emerg. Infect. Dis. 26 (7).

He, X., Lau, E.H., Wu, P., Deng, X., Wang, J., Hao, X., et al., 2020. Temporal dynamics in viral shedding and transmissibility of COVID-19. medRxiv doi:10.1038/s41591-020-0869-5.

Jacobs, J.L., Ohde, S., Takahashi, O., Tokuda, Y., Omata, F., Fukui, T., 2009. Use of surgical face masks to reduce the incidence of the common cold among health care workers in Japan: a randomized controlled trial. Am. J. Infect. Control 37 (5), 417–419.

Jefferson, T., Del Mar, C., Dooley, L., Ferroni, E., Al-Ansary, L.A., Bawazeer, G.A., et al., 2009. Physical interventions to interrupt or reduce the spread of respiratory viruses: systematic review. BMJ (Clin. Res. Ed.) 339, b3675.

Jefferson, T., Del Mar, C.B., Dooley, L., Ferroni, E., Al-Ansary, L.A., Bawazeer, G.A., et al., 2011. Physical interventions to interrupt or reduce the spread of respiratory viruses. Cochr. Database Syst. Rev. (7), CD006207.

Jefferson, T., Foxlee, R., Del Mar, C., Dooley, L., Ferroni, E., Hewak, B., et al., 2008. Physical interventions to interrupt or reduce the spread of respiratory viruses: systematic review. BMJ (Clin. Res. Ed.) 336 (7635), 77–80.

Johnson, D., Druce, J.D., Birch, C., Grayson, M.L., 2009. A quantitative assessment of the efficacy of surgical and N95 masks to filter influenza virus in patients with acute influenza infection. Clin. Infect. Dis. 49 (2), 275–277.

Keeley, A.J., Evans, C., Colton, H., Ankcorn, M., Cope, A., Bennett, T., et al., 2020. Roll-out of SARS-CoV-2 testing for healthcare workers at a large NHS Foundation Trust in the United Kingdom, March 2020. Eurosurveillance 25 (14), 2000433.

Larson, E.L., Ferng, Y.-.H., Wong-McLoughlin, J., Wang, S., Haber, M., Morse, S.S., 2010. Impact of non-pharmaceutical interventions on URIs and influenza in crowded, urban households. Public Health Rep. 125 (2), 178–191.

Lee K.M., Shukla V.K., Clark M., Mierzwinski-Urban M., Pessoa-Silva C.L., Conly J. Physical Interventions to Interrupt or Reduce the Spread of Respiratory Viruses - Resource Use Implications: a Systematic Review. Ottawa ON: 2011 Cadth.; 2011 Dec.

Leung, N.H., Chu, D.K., Shiu, E.Y., Chan, K.-.H., McDevitt, J.J., Hau, B.J., et al., 2020. Respiratory virus shedding in exhaled breath and efficacy of face masks. Nat. Med. 1–5.

Liu, Y., Ning, Z., Chen, Y., Guo, M., Liu, Y., Gali, N.K., et al., 2020. Aerodynamic characteristics and RNA concentration of SARS-CoV-2 aerosol in Wuhan hospitals during COVID-19 outbreak. bioRxiv doi:10.1101/2020.03.08.982637.

Loeb, M., Dafoe, N., Mahony, J., John, M., Sarabia, A., Glavin, V., et al., 2009. Surgical mask vs N95 respirator for preventing influenza among health care workers: a randomized trial. JAMA 302 (17), 1865–1871.

MacIntyre, C.R., Cauchemez, S., Dwyer, D.E., Seale, H., Cheung, P., Browne, G., et al., 2009. Face mask use and control of respiratory virus transmission in households. Emerg. Infect. Dis. 15 (2), 233.

MacIntyre, C.R., Chughtai, A.A., Seale, H., Richards, G.A., Davidson, P.M., 2014. Respiratory protection for healthcare workers treating Ebola virus disease (EVD): are facemasks sufficient to meet occupational health and safety obligations? Int. J. Nurs. Stud. 51 (11), 1421–1426.

MacIntyre, C.R., Chughtai, A.A., Seale, H., Richards, G.A., Davidson, P.M., 2014. Uncertainty, risk analysis and change for Ebola personal protective equipment guidelines. Int. J. Nurs. Stud. 52 (5), 899–903.

MacIntyre, C.R., Chughtai, A.A., Rahman, B., Peng, Y., Zhang, Y., Seale, H., et al., 2017. The efficacy of medical masks and respirators against respiratory infection in healthcare workers. Influenza Other Respir. Viruses 11 (6), 511–517.

Macintyre, R., Chughtai, A., Tham, C.D., Seale, H., 2020. Covid-19: Should cloth masks be used by healthcare workers as a last resort? BMJ (Clin. Res. Ed.) [Internet][cited 2020]. https://blogs.bmj.com/bmj/2020/04/09/covid-19-should-cloth-masks-be-used-by-healthcare-workers-as-a-last-resort/.

MacIntyre, C.R., Seale, H., Dung, T.C., Hien, N.T., Nga, P.T., Chughtai, A.A., et al., 2015. A cluster randomised trial of cloth masks compared with medical masks in healthcare workers. BMJ Open 5 (4), e006577.

MacIntyre, C.R., Wang, Q., Cauchemez, S., Seale, H., Dwyer, D.E., Yang, P., et al., 2011. A cluster randomized clinical trial comparing fit-tested and non-fit-tested N95 respirators to medical masks to prevent respiratory virus infection in health care workers. Influenza Other Respir. Viruses 5 (3), 170–179.

MacIntyre, C.R., Wang, Q., Seale, H., Yang, P., Shi, W., Gao, Z., et al., 2013. A randomized clinical trial of three options for N95 respirators and medical masks in health workers. Am. J. Respir. Crit. Care Med. 187 (9), 960–966.

MacIntyre, C.R., Zhang, Y., Chughtai, A.A., Seale, H., Zhang, D., Chu, Y., et al., 2016. Cluster randomised controlled trial to examine medical mask use as source control for people with respiratory illness. BMJ Open 6 (12), e012330.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., et al., 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst. Rev. 4 (1), 1.

Ong, S.W.X., Tan, Y.K., Chia, P.Y., Lee, T.H., Ng, O.T., Wong, M.S.Y., et al., 2020. Air, surface environmental, and personal protective equipment contamination by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) from a symptomatic patient. JAMA: J. Am. Med. Assoc. doi:10.1001/jama.2020.3227.

Radonovich, L.J., Simberkoff, M.S., Bessesen, M.T., Brown, A.C., Cummings, D.A.T., Gaydos, C.A., et al., 2019. N95 respirators vs medical masks for preventing influenza among health care personnel: a randomized clinical trial. JAMA J. Am. Med. Assoc. 322 (9), 824–833.

Salgado, C.D., Farr, B.M., Hall, K.K., Hayden, F.G., 2002. Influenza in the acute hospital setting. Lancet Infect. Dis. 2 (3), 145–155.

Santarpia, J.L., Rivera, D.N., Herrera, V., Morwitzer, M.J., Creager, H., Santarpia, G.W., et al., medRxiv. 2020. Transmission Potential of SARS-CoV-2 in Viral Shedding Observed at the. University of Nebraska Medical Center doi:10.1101/2020.03.23.20039446.

Simmerman, J.M., Suntarattiwong, P., Levy, J., Jarman, R.G., Kaewchana, S., Gibbons, R.V., et al., 2011. Findings from a household randomized controlled trial of hand washing and face masks to reduce influenza transmission in Bangkok, Thailand. Influenza Other Respir. Viruses 5 (4), 256–267.

Suess, T., Remschmidt, C., Schink, S.B., Schweiger, B., Nitsche, A., Schroeder, K., et al., 2012. The role of facemasks and hand hygiene in the prevention of influenza transmission in households: results from a cluster randomised trial; Berlin, Germany, 2009-2011. BMC Infect. Dis. 12, 26.

van Doremalen, N., Bushmaker, T., Morris, D.H., Holbrook, M.G., Gamble, A., Williamson, B.N., et al., 2020. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. New Engl. J. Med. 382 (16), 1564–1567. doi:10.1056/NEJMc2004973.

Verbeek, J.H., Rajamaki, B., Ijaz, S., Sauni, R., Toomey, E., Blackwood, B., et al., 2020. Personal protective equipment for preventing highly infectious diseases due to exposure to contaminated body fluids in healthcare staff. Cochr. Database Syst. Rev. (4).

World Health Organisation (WHO). Infection prevention and control during health care when novel coronavirus (nCoV) infection is suspected. Interim guidance 2020 [Available from: https://www.who.int/publications-detail/infection-prevention-and-control-during-health-care-when-novel-coronavirus-(ncov)-infection-is-suspected.

Zhan, M., Qin, Y., Xue, X., Zhu, S., 2020. Death from Covid-19 of 23 Health Care Workers in China. New Engl. J. Med. doi:10.1056/NEJMc2005696.

Zhang, W., Du, R.-.H., Li, B., Zheng, X.-.S., Yang, X.-.L., Hu, B., et al., 2020. Molecular and serological investigation of 2019-nCoV infected patients: implication of multiple shedding routes. Emerg. Microbes Infect. 9 (1), 386–389.

Zou, L., Ruan, F., Huang, M., Liang, L., Huang, H., Hong, Z., et al., 2020. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. New Engl. J. Med. 382 (12), 1177–1179. doi:10.1056/NEJMc2001737.

ScienceDaily®
Your source for the latest research news

A replaceable, more efficient filter for N95 masks

Date: May 21, 2020

Source: American Chemical Society

Summary: Researchers have developed a membrane that can be attached to a regular N95 mask and replaced when needed. The filter has a smaller pore size than normal N95 masks, potentially blocking more virus particles.

FULL STORY

Since the outbreak of COVID-19, there's been a worldwide shortage of face masks -- particularly, the N95 ones worn by health care workers. Although these coverings provide the highest level of protection currently available, they have limitations. Now, researchers reporting in *ACS Nano* have developed a membrane that can be attached to a regular N95 mask and replaced when needed. The filter has a smaller pore size than normal N95 masks, potentially blocking more virus particles.

N95 masks filter about 85% of particles smaller than 300 nm. SARS-CoV-2 (the coronavirus that causes COVID-19) is in the size range of 65-125 nm, so some virus particles could slip through these coverings. Also, because of shortages, many health care workers have had to wear the same N95 mask repeatedly, even though they are intended for a single use. To help overcome these problems, Muhammad Mustafa Hussain and colleagues wanted to develop a membrane that more efficiently filters particles the size of SARS-CoV-2 and could be replaced on an N95 mask after every use.

To make the membrane, the researchers first developed a silicon-based, porous template using lithography and chemical etching. They placed the template over a polyimide film and used a process called reactive ion etching to make pores in the membrane, with sizes ranging from 5-55 nm. Then, they peeled off the membrane, which could be attached to an N95 mask. To ensure that the nanoporous membrane was breathable, the researchers measured the airflow rate through the pores. They found that for pores tinier than 60 nm (in other words, smaller than SARS-CoV-2), the pores needed to be placed a maximum of 330 nm from each other to achieve good breathability. The hydrophobic membrane also cleans itself because droplets slide off it, preventing the pores from getting clogged with viruses and other particles.

The authors acknowledge funding from the King Abdullah University of Science and Technology Office of Sponsored Research.

Story Source:

Materials provided by **American Chemical Society**. *Note: Content may be edited for style and length.*

Journal Reference:

1. Nazek El-Atab, Nadeem Qaiser, Huda Badghaish, Sohail F. Shaikh, Muhammad Mustafa Hussain. **Flexible Nanoporous Template for the Design and Development of Reusable Anti-COVID-19 Hydrophobic Face Masks**. *ACS Nano*, 2020; DOI: 10.1021/acsnano.0c03976

Cite This Page:

MLA APA Chicago

American Chemical Society. "A replaceable, more efficient filter for N95 masks." ScienceDaily. ScienceDaily, 21 May 2020. <www.sciencedaily.com/releases/2020/05/200521124646.htm>.

RELATED STORIES

Inexpensive, Accessible Device Provides Visual Proof That Masks Block Droplets

Aug. 11, 2020 — In a proof-of-concept study, researchers report that a simple, low-cost technique provided visual proof that face masks are effective in reducing droplet emissions during normal wear. They found that ...

Elastomeric Masks Provide a More Durable, Less Costly Option for Health Care Workers

June 12, 2020 — A cost-effective strategy for health care systems to offset N95 mask shortages due to COVID-19 is to switch to reusable elastomeric respirator ...

Heating Could Be the Best Way to Disinfect N95 Masks for Reuse

May 5, 2020 — Since the outbreak of the COVID-19 pandemic, N95 face masks have been in short supply. Health care workers, in particular, desperately need these masks to protect themselves from the respiratory ...

Surgeons Help Create New Process for Disinfecting and Reusing N95 Masks

Apr. 28, 2020 — Amid shortages of personal protective equipment due to the Coronavirus Disease 2019 (COVID-19) pandemic, a St. Louis health care system has implemented a process to disinfect disposable N95 ...

Surgical Mask vs N95 Respirator for Preventing Influenza Among Health Care Workers

A Randomized Trial

Mark Loeb, MD, MSc
Nancy Dafoe, RN
James Mahony, PhD
Michael John, MD
Alicia Sarabia, MD
Verne Glavin, MD
Richard Webby, PhD
Marek Smieja, MD
David J. D. Earn, PhD
Sylvia Chong, BSc
Ashley Webb, BS
Stephen D. Walter, PhD

Context Data about the effectiveness of the surgical mask compared with the N95 respirator for protecting health care workers against influenza are sparse. Given the likelihood that N95 respirators will be in short supply during a pandemic and not available in many countries, knowing the effectiveness of the surgical mask is of public health importance.

Objective To compare the surgical mask with the N95 respirator in protecting health care workers against influenza.

Design, Setting, and Participants Noninferiority randomized controlled trial of 446 nurses in emergency departments, medical units, and pediatric units in 8 tertiary care Ontario hospitals.

Intervention Assignment to either a fit-tested N95 respirator or a surgical mask when providing care to patients with febrile respiratory illness during the 2008-2009 influenza season.

Main Outcome Measures The primary outcome was laboratory-confirmed influenza measured by polymerase chain reaction or a 4-fold rise in hemagglutinin titers. Effectiveness of the surgical mask was assessed as noninferiority of the surgical mask compared with the N95 respirator. The criterion for noninferiority was met if the lower limit of the 95% confidence interval (CI) for the reduction in incidence (N95 respirator minus surgical group) was greater than –9%.

Results Between September 23, 2008, and December 8, 2008, 478 nurses were assessed for eligibility and 446 nurses were enrolled and randomly assigned the intervention; 225 were allocated to receive surgical masks and 221 to N95 respirators. Influenza infection occurred in 50 nurses (23.6%) in the surgical mask group and in 48 (22.9%) in the N95 respirator group (absolute risk difference, –0.73%; 95% CI, –8.8% to 7.3%; $P = .86$), the lower confidence limit being inside the noninferiority limit of –9%.

Conclusion Among nurses in Ontario tertiary care hospitals, use of a surgical mask compared with an N95 respirator resulted in noninferior rates of laboratory-confirmed influenza.

Trial Registration clinicaltrials.gov Identifier: NCT00756574

JAMA. 2009;302(17):1865-1871 www.jama.com

INFLUENZA CAUSES ANNUAL EPIDEMICS of respiratory illness worldwide and is the most important cause of medically attended acute respiratory illness.[1,2] Moreover, there is increasing concern about the recently declared influenza pandemic due to 2009 influenza A(H1N1) in humans.[3-5]

Transmission of influenza can occur by coughing or sneezing where infectious particles of variable size, ranging from approximately 0.1 to 100 µm, may be inhaled.[6] This range of particles has a yet undefined but possibly important role in transmission. Although data from animal models and human experimental studies suggest that short-range inhalational transmission with small droplet nuclei (<10 µm) can occur,[7-11] the exact nature of transmission of influenza that occurs

For editorial comment see p 1903.

Author Affiliations: Departments of Pathology and Molecular Medicine (Drs Loeb, Mahony, and Smieja and Ms Dafoe), Medicine (Dr Loeb), Clinical Epidemiology and Biostatistics (Drs Loeb, Smieja, Earn, and Walter), and Mathematics and Statistics (Dr Earn), Michael G. DeGroote Institute for Infectious Disease Research (Drs Loeb, Mahony, Smieja, and Earn), McMaster University, Hamilton, Ontario, Canada; St Joseph's Hospital Regional Virology Laboratory, Hamilton (Dr Mahony and Ms Chong); Departments of Pathology and Microbiology and Immunology, University of Western Ontario, London (Dr John); Department of Microbiology, Credit Valley Hospital, Mississauga, Ontario (Dr Sarabia); Joseph Brant Memorial Hospital, Burlington, Ontario (Dr Glavin); and World Health Organization Collaborating Center for Studies on the Ecology of Influenza in Animals and Birds, St Judes Children's Hospital, Memphis, Tennessee (Dr Webby and Ms Webb).
Corresponding Author: Mark Loeb, MD, MSc, McMaster University, 1200 Main St W, Hamilton, ON L8N 3Z5, Canada (loebm@mcmaster.ca).

in nonexperimental settings is not well understood.[12] As a consequence, considerable uncertainty exists about the effectiveness of personal respiratory devices against influenza for health care workers.

During a pandemic, reducing transmission of influenza to health care workers may not only help support the health care workforce, but may also prevent influenza transmission to patients. Other personal protective strategies, such as effective vaccines or antiviral drugs, may be limited in availability. Given the likelihood that N95 respirators will be in short supply during a pandemic and unavailable in many countries, understanding the relative effectiveness of personal respiratory protective equipment is important. There are few comparative studies of respiratory protective devices,[13-15] and data comparing the surgical mask with the N95 respirator among health care workers are sparse.

We conducted a randomized trial to compare the surgical mask with the N95 respirator in health care workers. We hypothesized that the surgical mask, which is less expensive and more widely available than the N95 respirator, offers similar protection to the N95 respirator among health care workers at highest risk for exposure to influenza.

METHODS

Participants

We enrolled nurses who worked in emergency departments, medical units, and pediatric units in 8 Ontario tertiary care hospitals, of which 6 were within the greater Toronto area. Six of the 8 hospitals were university-affiliated teaching hospitals (range of bed size, 310-400) and 2 were community hospitals (bed sizes, 256 and 400). Participants were enrolled from a total of 22 units, which included 9 acute medical units, 7 emergency departments, and 6 pediatric units. There were an average of 34 beds (range, 14-60 beds) on the medical units and an average of 27 beds (range, 19-38) on the pediatric units.

Nurses expected to work full-time (defined as >37 hours per week) on study units during the 2008-2009 influenza season were eligible. Nurses had to provide current fit-test certification. Nurses who could not pass a fit test were excluded from the study. The research protocol was approved by the McMaster University research ethics review board. All participants gave written informed consent.

Interventions

Randomization was performed centrally by an independent clinical trials coordinating group such that investigators were blind to the randomization procedure and group assignment and was stratified by center in permuted blocks of 4 participants. It was not possible to conceal the identity of the N95 respirator or the surgical mask since manipulating these devices would interfere with their function. Laboratory personnel conducting hemagglutinin inhibition assays, polymerase chain reaction (PCR), and viral culture for influenza were blinded to allocation. Nurses allocated to the surgical mask group were required to wear the brand of surgical mask already in use at their hospital. Following the severe acute respiratory syndrome (SARS) outbreak in Ontario, use of such a surgical mask was required by the Ministry of Health and Long-Term Care when providing care to or when within 1 m of a patient with febrile respiratory illness, defined as symptoms of a body temperature 38°C or greater and new or worsening cough or shortness of breath.[16] Nurses were instructed in proper placement of the surgical mask according to the manufacturer's recommendations.

Since fit testing is mandatory for nurses in Ontario, the majority of nurses in the study had been fit tested prior to enrollment; additional fit testing was conducted for nurses who had not been fit tested in 2008. Using a standard protocol, a technician showed the participant how to position the respirator and fasten the strap and determine whether it provided an acceptable fit. The nurse was asked to wear the most comfortable mask for at least 5 minutes to assess fit. Adequacy of the respiratory fit was assessed using standard criteria, including chin placement, adequate strap tension, appropriate respirator size, fit across nose bridge, tendency of respirator to slip, and position of mask on face and cheeks. The nurse then conducted a user seal check.[17] Nurses had a qualitative fit testing using the saccharin or Bitrex protocol.[17]

Nurses were asked to begin using the surgical mask or N95 respirator when caring for patients with febrile respiratory illness at the beginning of the influenza season, which was defined as 2 or more consecutive isolations of influenza per week in each study region. Nurses wore gloves and gowns when entering the room of a patient with febrile respiratory illness, which was routine practice. For aerosol-generating procedures (such as intubation or bronchoscopy), as long as tuberculosis was not suspected, nurses continued to use the respiratory device they were assigned to.

We had planned to stop the study at the end of influenza season. However, because of the 2009 influenza A(H1N1) pandemic, the study was stopped on April 23, 2009, when the Ontario Ministry of Health and Long-Term Care recommended N95 respirators for all health care workers taking care of patients with febrile respiratory illness.

Follow-up

All participants were assessed for signs and symptoms of influenza twice weekly using Web-based questionnaires. Response to the questionnaire was monitored centrally and participants who failed to provide a response were contacted and asked to complete the questionnaire. If a new symptom was reported, the study nurse was notified and a flocked nasal specimen (Copan Italia, Brescia, Italy) was obtained by the participants. They were trained to insert the swab into the left or right nostril and rotate the swab at least 3 times and to conduct self-swabbing if

any of 1 of the following symptoms or signs were present: fever (temperature ≥38°C), cough, nasal congestion, sore throat, headache, sinus problems, muscle aches, fatigue, earache, ear infection, or chills. We also provided participants with tympanic thermometers. To assess household exposures between study groups, we asked participants whether household members (spouses, roommates, or children) had experienced influenza-like illness over the study period.

Outcomes

The primary outcome of this study was laboratory-confirmed influenza. This was defined by either the detection of viral RNA using reverse-transcriptase (RT) PCR from nasopharyngeal and flocked nasal specimens or at least a 4-fold rise in serum antibodies to circulating influenza strain antigens. All nasopharyngeal or nasal specimens were tested for influenza and other respiratory viruses with the xTAG Respiratory Virus Panel test (Luminex Molecular Diagnostics, Toronto, Ontario, Canada).[18] This multiplex PCR assay detects influenza A virus subtypes H1 (seasonal), H3, and H5 as well as the majority of other viruses that cause respiratory illness in humans.

Blood specimens for serology were obtained prior to enrollment and at the end of the follow-up period. Serological infection was defined by detection of 4-fold or greater increase in influenza-specific hemagglutinin inhibition assay titer between baseline and convalescent serum samples using guinea pig erythrocytes and the antigens circulating A/Brisbane/59/2007(H1N1)-like virus; A/Brisbane/10/2007(H3N2)-like virus; B/Florida/4/2006-like virus; and A/TN/1560/09(H1N1), the circulating pandemic influenza virus. For A/Brisbane/59/2007(H1N1)-like virus, A/Brisbane/10/2007(H3N2)-like virus, and B/Florida/4/2006-like virus, we restricted serological criteria of infection to nurses who did not receive the trivalent 2008-2009 influenza vaccine to reduce misclassification due to vaccine response.

Secondary outcomes included detection of the following noninfluenza viruses by PCR: parainfluenza virus types 1, 2, 3, and 4; respiratory syncytial virus types A and B; adenovirus; metapneumovirus; rhinovirus-enterovirus; and coronaviruses OC43, 229E, SARS, NL63, and HKU1. Influenza-like illness was defined as the presence of cough and fever (temperature ≥38°C).[19] Work-related absenteeism and physician visits for respiratory illness were also assessed.

Audits

To assess compliance of participants with the assigned mask or N95 respirator, we conducted audits during what we anticipated was peak influenza period, from March 11 to April 3, 2009. Medical and pediatric hospital study units at all centers with nurses participating in the study were contacted by telephone daily by a research assistant to assess whether there were patients admitted to the unit in droplet precautions for influenza or febrile respiratory illness. If there were such cases and if the primary nurse for the patient was enrolled in our study, a trained auditor was sent to the unit to observe for compliance. The auditor was instructed to stand a short distance from the patient isolation room to remain inconspicuous but within distance to accurately record the audit. Auditors were asked to remain on the unit until they recorded the type of protective equipment worn by the participant prior to the participant entering the isolation room.

To maintain patient confidentiality and to remain anonymous to the study participant, no audits were conducted within the patient's room. Once an audit was conducted, the session was completed. Audits were conducted both on weekdays and on weekends during day and evening shifts. Assessment of hand hygiene was not conducted.

Statistical Analysis

The effectiveness of the surgical mask was assessed through a noninferiority analysis relative to the N95 respirator.[20] For the primary analysis, the difference in the incidence of laboratory-confirmed influenza between the N95 respirator group and surgical mask group was estimated and the corresponding 2-sided 95% confidence interval (CI) was calculated. We used the Fisher exact test to assess statistical significance in contingency tables having expected cell frequencies less than 5. Noninferiority to the N95 respirator was achieved if the lower limit of the 95% CI for the reduction in incidence (N95 respirator minus surgical group) was greater than the prespecified noninferiority limit of –9%. Assuming an event rate of 20% in controls, this limit was selected on a clinical basis considering that laboratory-confirmed influenza would include asymptomatic cases in addition to symptomatic cases of influenza. Infection detected by serology can account for up to 75% of cases of laboratory-confirmed influenza where febrile illness is not present.[21]

Since we did not anticipate severe outcomes (eg, mortality) in the study sample, we used a similar approach for influenza-like illness, work-related absenteeism, and physician visits for respiratory illness. All participants who had follow-up data collected (ie, had not withdrawn prior to any follow-up after they had been randomized) were included in the analysis. Since intention-to-treat analyses in noninferiority trials may be biased toward finding no difference, we also conducted an analysis of our primary outcome using only data from participants with complete follow-up.[22]

To avoid lack of independence associated with counting multiple outcomes, each specific outcome in a participant was only counted once. With a power of 90% and a 2-sided type-I error rate of 5%, the required sample would be 191 participants in each group for a noninferiority test assuming an absolute risk reduction of 12% in the N95 respirator group compared with the surgical mask. If the absolute reduction was assumed to be 10%, a statistical power of 80% would be maintained. The absolute risk reductions selected

were based on consensus by clinician investigators. Assuming a 10% dropout rate, we estimated that a total of 420 participants would be needed. SAS version 9.1.3 (SAS Institute, Cary, North Carolina) was used to conduct the analyses.

RESULTS

Between September 23, 2008, and December 8, 2008, 478 nurses were assessed for eligibility and 446 participants from 8 centers in Ontario were enrolled. They were then randomly assigned the intervention, 225 to the surgical mask and 221 to the N95 respirator (FIGURE). The mean age of participants was 36.2 years, 94% of them were female, and study groups were well balanced in terms of demographics (TABLE 1). Vaccination status was similar: 68 participants (30.2%) in the surgical mask group and 62 (28.1%) in the N95 respirator group had received 2008-2009 trivalent inactivated influenza vaccine.

Figure. Flow Diagram for Trial of Surgical Mask vs N95 Respirator

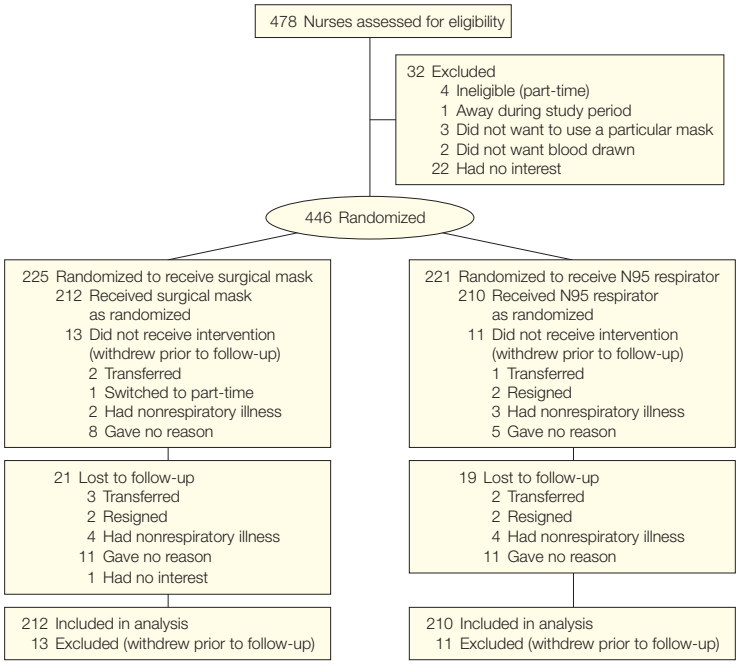


Table 1. Characteristics of 446 Nurse Participants in the Surgical Mask and N95 Respirator Groups

Characteristic	No. (%) Surgical Mask (n = 225)	No. (%) N95 Respirator (n = 221)
Age, mean (SD) [range], y	36.5 (10.6) [21-62]	35.8 (10.6) [21-60]
Female sex	212 (94.2)	208 (94.1)
Vaccinated against influenza	68 (30.2)	62 (28.1)
≥1 Coexisting conditions	22 (9.8)	26 (11.8)
Asthma	10 (4.4)	12 (5.4)
Diabetes	3 (1.3)	6 (2.7)
Metabolic	2 (1.0)	4 (1.8)
Immunocompromised[a]	3 (1.3)	3 (1.3)
Pregnancy	5 (2.2)	2 (0.9)
Other[b]	6 (2.7)	3 (1.3)
Distribution by hospital unit		
Medical	55 (24.4)	52 (23.5)
Pediatric	58 (26.2)	62 (28.1)
Emergency	112 (49.8)	107 (48.4)

[a] Immunosuppressive medications for transplantation (n=1), rheumatoid arthritis (n=3), uveitis (n=1), and Crohn disease (n=1).
[b] Includes chronic renal failure (n=1), coronary artery disease (n=1), liver disease (n=2), seizures/brain disorder (n=2), and connective tissue disease (n=4).

Follow-up began January 12, 2009, and ended April 23, 2009. Mean (SD) duration of follow-up was similar between groups: 97.9 (16.1) days in the surgical group and 97.2 (18.0) days in the N95 respirator group. There were 24 participants who withdrew from the study with no follow-up—13 in the surgical mask group and 11 in the N95 respirator group—because of resignation or transfer (n=5), working part-time (n=1), no response (n=13), or illness (n=5) (Figure). None of the health care workers withdrew because of respiratory illness. Of the resulting 422 (all of whom were in the analysis), follow-up was complete in 386 (91.4%), and 403 (95.5%) had acute and convalescent sera collected. There were 223 nasal specimens obtained (115 in the surgical mask group and 108 in the N95 respirator group).

Laboratory-confirmed influenza (by RT-PCR or ≥4-fold rise in serum titers) occurred in 50 nurses (23.6%) in the surgical mask group and in 48 (22.9%) in the N95 respirator group (absolute risk difference, –0.73%; 95% CI, –8.8% to 7.3%; *P*=.86), indicating noninferiority of the surgical mask (TABLE 2). The diagnosis of influenza was made by RT-PCR in 6 nurses (2.8%) in the surgical mask group (5 influenza A and 1 influenza B) and 4 (1.8%) in the N95 respirator group (1 influenza A and 3 influenza B) (absolute risk difference, –0.93%; 95% CI, –3.82% to 1.97%; *P*=.75). Four of the influenza A cases detected by PCR were H1 (all in the surgical mask group). The serology results are summarized in Table 2. Notably, 8.0% in the surgical mask group and 11.9% in the N95 respirator group had a

4-fold or greater rise in serum titers to A/TN/1560/09(H1N1), the circulating pandemic swine influenza strain. Noninferiority was demonstrated between the surgical mask group and the N95 respirator group for 2009 influenza A(H1N1) (absolute risk difference, 3.89%; 95% CI, –1.82% to 9.59%; $P=.18$).

When the analysis was conducted using only the data from participants with complete follow-up visits, laboratory-confirmed influenza (by RT-PCR or ≥4-fold rise in serum titers) occurred in 66 nurses (33.9%) in the surgical mask group and in 72 (37.7%) in the N95 respirator group (absolute risk difference, 3.85%; 95% CI, –5.71% to 13.41%; $P=.43$), indicating noninferiority.

No adenoviruses; no respiratory syncytial virus type A; and no parainfluenza 1, 2, and 4 viruses were detected by PCR. There were no significant differences between the surgical mask and N95 respirator groups in respiratory syncytial virus type B, metapneumovirus, parainfluenza 3, rhinovirus-enterovirus, or coronoviruses. The lower CIs for the differences were greater than –9%, meeting our criteria for noninferiority (TABLE 3). All 52 (100%) of those having infection with a respiratory virus other than influenza had 1 or more symptoms, but they did not meet the influenza-like illness definition.

Nine nurses (4.2%) in the surgical mask group and 2 nurses (1.0%) in the N95 respirator group met our criteria for influenza-like illness (absolute risk difference, –3.29%; 95% CI, –6.31% to 0.28%; $P=.06$) (TABLE 4). All 11 had laboratory-confirmed influenza. A significantly greater number of nurses in the surgical mask group (12, or 5.66%) reported fever compared with the N95 respirator group (2, or 0.9%; $P=.007$). There was no significant difference in nurses who reported cough, nasal congestion, headache, sore throat, myalgia, fatigue, earache, or ear infection. Of the 44 nurses in each group who had influenza diagnosed by serology, 29 (65.9%) in the surgical mask group and 31 (70.5%) in the N95 respirator group had no symptoms.

There were 13 physician visits (6.1%) for respiratory illness among those in the surgical mask group compared with 13 (6.2%) in the N95 respirator group (absolute risk difference, –0.06%; 95% CI, –4.53% to 4.65%; $P=.98$). Forty-two participants (19.8%) in the surgical mask group reported an episode of work-related absenteeism compared with 39 (18.6%) in the N95 respiratory group (absolute risk difference, –1.24%; 95% CI, –8.75% to 6.27%; $P=.75$) (Table 4). There were no episodes of lower respiratory tract infec-

Table 2. Comparison of Laboratory-Confirmed Influenza Between the Surgical Mask and N95 Respirator Groups

	No. (%)			
	Surgical Mask (n = 212)	N95 Respirator (n = 210)	Absolute Risk Difference, % (95% CI)	*P* Value
Laboratory-confirmed influenza[a]	50 (23.6)	48 (22.9)	–0.73 (–8.8 to 7.3)	.86
RT-PCR influenza A	5 (2.4)	1 (0.5)	–1.88 (–4.13 to 0.36)	.22
RT-PCR influenza B	1 (0.5)	3 (1.4)	0.96 (–0.89 to 2.81)	.37
≥4-Fold rise in serum titers A/Brisbane/59/2007 (H1N1)[b]	25 (11.8)	21 (10)	–1.79 (–7.73 to 4.15)	.55
≥4-Fold rise in serum titers A/Brisbane/10/2007 (H3N2)[b]	42 (19.8)	49 (23.3)	3.52 (–4.32 to 11.36)	.38
≥4-Fold rise in serum titers B/Florida/4/2006[b]	15 (7.1)	19 (9.0)	2.0 (–3.0 to 7.17)	.46
≥4-Fold rise in serum titers A/TN/1560/09 (H1N1)[b]	17 (8.0)	25 (11.9)	3.89 (–1.82 to 9.59)	.18

Abbreviations: CI, confidence interval; RT-PCR, reverse-transcriptase polymerase chain reaction.
[a] Influenza detected by 1 or more of the following: RT-PCR A, RT-PCR B, and ≥4-fold rise in serum titers to A/Brisbane/59/2007(H1N1), A/Brisbane/10/2007(H3N2), and B/Florida/4/2006. Serology includes only nonvaccinated nurses.
[b] Includes both vaccinated and nonvaccinated nurses. Two hundred ninety-four nurses were not vaccinated (147 in each group).

Table 3. Comparison of RT-PCR Results for Other Respiratory Viruses Between the Surgical Mask and N95 Respirator Groups

	No. (%)			
	Surgical Mask (n = 212)	N95 Respirator (n = 210)	Absolute Risk Difference, % (95% CI)	*P* Value
Respiratory syncytial virus[a]	2 (0.9)	1 (0.5)	–0.47 (–2.07 to 1.13)	>.99
Metapneumovirus	4 (1.9)	3 (1.4)	–0.46 (–1.98 to 2.89)	>.99
Parainfluenza virus[b]	1 (0.5)	2 (1.0)	0.48 (–1.12 to 2.09)	.62
Rhinovirus-enterovirus	8 (3.8)	10 (4.8)	0.99 (–2.87 to 4.85)	.62
Coronavirus[c]	9 (4.3)	12 (5.7)	1.47 (–2.68 to 5.62)	.49
Total[d]	20 (9.4)	22 (10.5)	1.04 (–4.67 to 6.76)	.72

Abbreviations: CI, confidence interval; RT-PCR, reverse-transcriptase polymerase chain reaction.
[a] Refers to respiratory syncytial virus type B only because no type A was detected.
[b] Refers to parainfluenza 3 only because no parainfluenza 1, 2, or 4 was detected.
[c] Refers to coronaviruses OC43, 229E, NL63, and HKU1.
[d] Totals are less than sums because more than 1 virus was detected in some participants.

Table 4. Clinical Outcomes Between the Surgical Mask and N95 Respirator Groups

	No. (%)			
	Surgical Mask (n = 212)	N95 Respirator (n = 210)	Absolute Risk Difference, % (95% CI)	*P* Value
Physician visits for respiratory illness	13 (6.1)	13 (6.2)	–0.06 (–4.53 to 4.65)	.98
Influenza-like illness[a]	9 (4.2)	2 (1.0)	–3.29 (–6.31 to 0.28)	.06
Work-related absenteeism	42 (19.8)	39 (18.6)	–1.24 (–8.75 to 6.27)	.75

Abbreviation: CI, confidence interval.
[a] Influenza-like illness was defined as the presence of both cough and temperature 38°C or greater.

tion among participants. There were no adverse events reported by participants.

Fifty-five participants (25.9%) in the surgical mask group vs 47 (22.4%) in the N95 respirator group reported a spouse or roommate with influenza-like illness ($P=.39$). Forty-eight participants (22.6%) in the surgical mask group vs 43 (20.5%) in the N95 respirator group reported a child with influenza-like illness ($P=.59$).

Over the 2-week audit period, there were 18 episodes of patients admitted to units in droplet precautions for influenza or febrile respiratory illness where the nurse providing care for the patient had been enrolled in our study. The results of the audit demonstrated that all 11 participants (100%) allocated to surgical masks and 6 of 7 participants (85.7%) allocated to N95 respirators were wearing the device to which they had been assigned.

COMMENT

Our data show that the incidence of laboratory-confirmed influenza was similar in nurses wearing the surgical mask and those wearing the N95 respirator. Surgical masks had an estimated efficacy within 1% of N95 respirators. Based on the prespecified definition, the lower CI for the difference in effectiveness of the surgical mask and N95 mask was within –9% and the statistical criterion of noninferiority was met. That is, surgical masks appeared to be no worse, within a prespecified margin, than N95 respirators in preventing influenza.

Transmission by small droplet spread would be compatible with greater protection with the N95 mask compared with the surgical mask where efficiency estimates range from 2% to 92% for particles smaller than 20 µm in diameter.[23-28] The fact that attack rates were similar may suggest that small aerosols did not dominate transmission.

One frequently cited concern about the surgical mask is its inability to obtain an appropriate seal compared with the N95 respirator.[29] Based on the results of this trial, this concern does not seem to be associated with an increased rate of infection of influenza or other respiratory viruses.

Influenza attack rates among health care workers in non-outbreak settings are sparse. Our data provide estimates of an attack rate (23%) in a largely unvaccinated cohort of nurses followed closely during a period of relatively mild influenza-like illness and into the beginning of what is now considered a pandemic period. Given that serology captures exposure over the entire season and that nurses have repeated exposures, this rate of infection was not unexpected. Our serological data in unvaccinated nurses were 20% for H3N2, 10% for H1N1, and 8% for influenza B. In a community-based study, age-specific rates of infection for those aged 30 to 39 years by serology was 16% for H3N2, approximately 5% for H1N1, and 5% for influenza B.[21] It is for this reason that the number of participants with influenza-like illness, defined by fever and cough alone,[19] were relatively few compared with the number with laboratory-confirmed influenza. Given that there was no difference in laboratory-confirmed influenza between study groups, the higher proportion of nurses in the surgical mask group with influenza-like illness, although not statistically significant, was unexpected.

The results of seroconversion to 2009 influenza A(H1N1) (10%) was unexpected given that the convalescent specimens were obtained from April 23 to May 15, 2009. This attack rate may suggest that 2009 influenza A(H1N1) was circulating in Ontario before April 2009. An alternative explanation for this high rate of seroconversion may be cross-reaction due to exposure to seasonal H1N1.

Strengths of this study include individual-level randomization, comprehensive laboratory-confirmed outcome assessment with PCR and serological evaluation, follow-up over an entire influenza season, and excellent participant follow-up.

There are a number of limitations of this study. Compliance with the intervention could not be assessed for all participants. Only 1 room entry was recorded per observation and the auditor did not enter the isolation room to assess whether the participant removed the respirator protection. Audits were only conducted on medical and pediatric units, not in the emergency department. Had there been poor compliance with the N95 respirator, this could have biased the study toward noninferiority. However, the results from our audited sample suggest excellent adherence. This is in keeping with the fact that all hospitals in the study were in Ontario, which was affected by the SARS outbreak and where use of personal protective equipment is mandated and audited by the Ontario Ministry of Labour.

We acknowledge that our protocol did not account for the effect of indirect contact because hand hygiene and use of gloves and gowns were not monitored. An imbalance in hand hygiene between study groups, with worse adherence in the N95 group, would have biased the study toward noninferiority. However, individual-level randomization and stratified randomization within hospitals would help balance any differences in adherence to hand hygiene between study groups. Because the use of gloves and gowns when entering the room of a patient with febrile respiratory illness was standard practice in our study hospitals, variability of use would likely have been minimal.

It is also impossible to determine whether participants acquired influenza due to hospital or community exposure. However, our data on household exposure suggest that such exposures were balanced between intervention groups. We acknowledge that not surveying participants' coworkers about influenza-like illness was a limitation. Since we did not collect information on droplet isolation precautions, a greater exposure of N95 respirator nurses vs surgical mask nurses to patients on droplet precautions would

have biased the study toward noninferiority. However, the fact that the nurses were well balanced on each ward and in the number of specimens obtained on each unit would minimize the chance of such differential exposure having occurred.

The major implication of this study is that protection with a surgical mask against influenza appears to be similar to the N95 respirator, meeting criteria for noninferiority. Our findings apply to routine care in the health care setting. They should not be generalized to settings where there is a high risk for aerosolization, such as intubation or bronchoscopy, where use of an N95 respirator would be prudent. In routine health care settings, particularly where the availability of N95 respirators is limited, surgical masks appear to be noninferior to N95 respirators for protecting health care workers against influenza.

Published Online: October 1, 2009 (doi:10.1001/jama.2009.1466).

Author Contributions: Dr Loeb had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Loeb, Webby, Smieja, Earn, Walter.

Acquisition of data: Loeb, Dafoe, Mahony, John, Sarabia, Glavin, Chong, Webb.

Analysis and interpretation of data: Loeb, Smieja, Chong, Walter.

Drafting of the manuscript: Loeb, Dafoe.

Critical revision of the manuscript for important intellectual content: Loeb, Mahony, John, Sarabia, Glavin, Webby, Smieja, Earn, Chong, Webb, Walter.

Statistical analysis: Loeb, Walter.

Obtained funding: Loeb, Walter.

Administrative, technical, or material support: Dafoe, John, Sarabia, Smieja, Earn, Chong, Webb.

Study supervision: Loeb, Mahony, Webby.

Financial Disclosures: None reported.

Funding/Support: This study was supported by the Public Health Agency of Canada.

Role of the Sponsor: The funding organization had no role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript.

REFERENCES

1. Langley JM, Faughnan ME. Prevention of influenza in the general population. *CMAJ*. 2004;171(10):1213-1222.

2. Thompson WW, Shay DK, Weintraub E, et al. Influenza-associated hospitalizations in the United States. *JAMA*. 2004;292(11):1333-1340.

3. Dawood FS, Jain S, Finelli L, et al; Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N Engl J Med*. 2009;360(25):2605-2615.

4. Cohen J, Enserink M. Swine flu: after delays, WHO agrees: the 2009 pandemic has begun. *Science*. 2009;324(5934):1496-1497.

5. Neumann G, Noda T, Kawaoka Y. Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature*. 2009;459(7249):931-939.

6. Nicas M, Nazaroff WW, Hubbard A. Toward understanding the risk of secondary airborne infection: emission of respirable pathogens. *J Occup Environ Hyg*. 2005;2(3):143-154.

7. Snyder MH, Stephenson EH, Young H, et al. Infectivity and antigenicity of live avian-human influenza A reassortment virus: comparison of intranasal and aerosol routes in squirrel monkeys. *J Infect Dis*. 1986;154(4):709-711.

8. Loosli CG, Lemon HM, Robertson OH, Appel E. Experimental airborne influenza infection: I, influence of humidity on survival of virus in air. *Proc Soc Exp Biol Med*. 1943;53:205-206.

9. Alford RH, Kasel JA, Gerone PJ, Knight V. Human influenza resulting from aerosol inhalation. *Proc Soc Exp Biol Med*. 1966;122(3):800-804.

10. Munster VJ, de Wit E, van den Brand JM, et al. Pathogenesis and transmission of swine-origin 2009 A(H1N1) influenza virus in ferrets. *Science*. 2009;325(5939):481-483.

11. Maines TR, Jayaraman A, Belser JA, et al. Transmission and pathogenesis of swine-origin 2009 A(H1N1) influenza viruses in ferrets and mice. *Science*. 2009;325(5939):484-487.

12. Brankston G, Gitterman L, Hirji Z, Lemieux C, Gardam M. Transmission of influenza A in human beings. *Lancet Infect Dis*. 2007;7(4):257-265.

13. Agah R, Cherry JD, Garakian AJ, Chapin M. Respiratory syncytial virus (RSV) infection rate in personnel caring for children with RSV infections: routine isolation procedure vs routine procedure supplemented by use of masks and goggles. *Am J Dis Child*. 1987;141(6):695-697.

14. MacIntyre CR, Cauchemez S, Dwyer DE, et al. Face mask use and control of respiratory virus transmission in households. *Emerg Infect Dis*. 2009;15(2):233-241.

15. Cowling BJ, Chan KH, Fang VJ, et al. Facemasks and hand hygiene to prevent influenza transmission in households: a randomized trial [published online August 3, 2009]. *Ann Intern Med*. http://www.annals.org/cgi/content/full/0000605-200910060-00142v1#FN.

16. Preventing respiratory illnesses: protecting patient and staff: infection control and surveillance standards for febrile respiratory illness (FRI) in non-outbreak conditions in acute care hospitals [September 2005]. Ministry of Health and Long-Term Care, Public Health Division, Provincial Infectious Diseases Advisory Committee. http://www.health.gov.on.ca/english/providers/program/infectious/diseases/best_prac/bp_fri_080406.pdf. Accessed September 11, 2009.

17. OSHA Technical Manual: section XIII: chapter 2: respiratory protection. US Department of Labor. http://www.osha.gov/dts/osta/otm/otm_viii/otm_viii_2.html. Accessed September 21, 2009.

18. Mahony J, Chong S, Merante F, et al. Development of a respiratory virus panel test for detection of twenty human respiratory viruses by use of multiplex PCR and a fluid microbead-based assay. *J Clin Microbiol*. 2007;45(9):2965-2970.

19. Call SA, Vollenweider MA, Hornung CA, Simel DL, McKinney WP. Does this patient have influenza? *JAMA*. 2005;293(8):987-997.

20. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA*. 2006;295(10):1152-1160.

21. Monto AS, Koopman JS, Longini IM Jr. Tecumseh study of illness: XIII, influenza, infection, and disease, 1976-1981. *Am J Epidemiol*. 1985;121(6):811-822.

22. Gotzsche PC. Lessons from and cautions about noninferiority and equivalence randomized trials. *JAMA*. 2006;295(10):1172-1174.

23. Brosseau LM, McCullough NV, Vesley D, Vincent JH. Mycobacterial aerosol collection efficiency of respirator and surgical mask filters under varying conditions of flow and humidity. *Appl Occup Environ Hyg*. 1997;12(6):435-445.

24. Cooper DW, Hinds WC, Price JM, Weker R, Yee HS. Common materials for emergency respiratory protection: leakage tests with a manikin. *Am Ind Hyg Assoc J*. 1983;44(10):720-726.

25. McCullough NV, Brosseau LM, Vesley D. Collection of three bacterial aerosols by respirator and surgical mask filters under varying conditions of flow and relative humidity. *Ann Occup Hyg*. 1997;41(6):677-690.

26. Mitakakis TZ, Tovey ER, Yates DH, et al. Particulate masks and non-powdered gloves reduce latex allergen inhaled by healthcare workers. *Clin Exp Allergy*. 2002;32(8):1166-1169.

27. Tuomi T. Face seal leakage of half masks and surgical masks. *Am Ind Hyg Assoc J*. 1985;46(6):308-312.

28. Willeke K, Qian YG, Donnelly J, Grinshpun S, Ulevicius V. Penetration of airborne microorganisms through a surgical mask and a dust/mist respirator. *Am Ind Hyg Assoc J*. 1996;57(4):348-355.

29. Influenza transmission and the role of personal protective respiratory equipment: an assessment of the evidence. Expert Panel on Influenza and Personal Protective Respiratory Equipment, Council of Canadian Academies. http://www.scienceadvice.ca/documents/(2007-12-19)_Influenza_PPRE_Final_Report.pdf. Accessed September 11, 2009.

Effectiveness of N95 respirators versus surgical masks in protecting health care workers from acute respiratory infection: a systematic review and meta-analysis

Jeffrey D. Smith MSc, Colin C. MacDougall MSc, Jennie Johnstone MD PhD, Ray A. Copes MD, Brian Schwartz MD, Gary E. Garber MD

Abstract

Background: Conflicting recommendations exist related to which facial protection should be used by health care workers to prevent transmission of acute respiratory infections, including pandemic influenza. We performed a systematic review of both clinical and surrogate exposure data comparing N95 respirators and surgical masks for the prevention of transmissible acute respiratory infections.

Methods: We searched various electronic databases and the grey literature for relevant studies published from January 1990 to December 2014. Randomized controlled trials (RCTs), cohort studies and case–control studies that included data on health care workers wearing N95 respirators and surgical masks to prevent acute respiratory infections were included in the meta-analysis. Surrogate exposure studies comparing N95 respirators and surgical masks using manikins or adult volunteers under simulated conditions were summarized separately. Outcomes from clinical studies were laboratory-confirmed respiratory infection, influenza-like illness and workplace absenteeism. Outcomes from surrogate exposure studies were filter penetration, face-seal leakage and total inward leakage.

Results: We identified 6 clinical studies (3 RCTs, 1 cohort study and 2 case–control studies) and 23 surrogate exposure studies. In the meta-analysis of the clinical studies, we found no significant difference between N95 respirators and surgical masks in associated risk of (a) laboratory-confirmed respiratory infection (RCTs: odds ratio [OR] 0.89, 95% confidence interval [CI] 0.64–1.24; cohort study: OR 0.43, 95% CI 0.03–6.41; case–control studies: OR 0.91, 95% CI 0.25–3.36); (b) influenza-like illness (RCTs: OR 0.51, 95% CI 0.19–1.41); or (c) reported workplace absenteeism (RCT: OR 0.92, 95% CI 0.57–1.50). In the surrogate exposure studies, N95 respirators were associated with less filter penetration, less face-seal leakage and less total inward leakage under laboratory experimental conditions, compared with surgical masks.

Interpretation: Although N95 respirators appeared to have a protective advantage over surgical masks in laboratory settings, our meta-analysis showed that there were insufficient data to determine definitively whether N95 respirators are superior to surgical masks in protecting health care workers against transmissible acute respiratory infections in clinical settings.

Competing interests: None declared.

This article has been peer reviewed.

Accepted: Dec. 2, 2015
Online: Mar. 7, 2016

Correspondence to:
Gary Garber, gary.garber@oahpp.ca

CMAJ **2016. DOI:10.1503/cmaj.150835**

Transmission of acute respiratory infections occurs primarily by contact and droplet routes, and accordingly, the use of a surgical mask, eye protection, gown and gloves should be considered appropriate personal protective equipment when providing routine care for a patient with a transmissible acute respiratory infection.[1–3] Concerns have been raised about possible acute respiratory infection spread via limited-distance airborne transmission, but this is controversial and has not been proven.[1,4–9] Also, experimental data suggest the superiority of N95 filtering facepiece respirators (N95 respirators) over surgical masks for the prevention of acute respiratory infections.[1] Randomized controlled trials (RCTs) and observational studies comparing N95 respirators and surgical masks have not shown a benefit, but they may have been underpowered.[10–17]

The lack of clarity has led to conflicting guideline recommendations regarding respiratory protective equipment for the prevention of acute respiratory infections: N95 respirators are recommended in some guidelines but not others.[18] Since the outbreak of severe acute respiratory syndrome (SARS), there has been a heightened level of controversy within Canada in determining the optimal ways to protect health care workers from respiratory pathogens. Conflicting recommendations from federal and provincial health authorities lead to confusion among heath care workers, which can result in lack of adherence to basic infection control principles and practices.

We performed a systematic review to assess and synthesize the available body of literature regarding N95 respirators versus surgical masks for the protection of health care workers against acute respiratory infections in a health care setting.

Methods

A detailed protocol developed a priori is described in Appendix 1 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.150835/-/DC1).

Literature search

We searched MEDLINE, Embase, the Database of Abstracts of Reviews of Effects, the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, Health Technology Assessment, the Collective Index of Nursing and Allied Health Literature, PsycINFO and Scopus for pertinent English-language studies published from Jan. 1, 1990, to Dec. 9, 2014. (The search strategies are available in Appendix 1, Tables S1–S9.) The search start date marks 4 years before N95 respirators became a part of standard respiratory protective equipment among health care workers in the United States.

We also conducted searches of the grey literature to obtain unpublished data. These searches were limited to the past 5 years (see Appendix 1, Table S10, for search details).

Study selection

Randomized controlled trials, prospective and retrospective cohort studies, and case–control studies were eligible for inclusion in the meta-analysis. Participants in clinical studies were health care workers in a health care setting. We defined health care worker as any worker in a health care setting who might be exposed to a patient with an acute respiratory infection. We excluded studies that solely involved protection of patients or community populations.

Surrogate exposure studies (i.e., experiments involving manikins or volunteers exposed to artificially produced aerosols) were not eligible for inclusion in the meta-analysis but were summarized to provide an overview of the laboratory-based experimental evidence for use of N95 respirators to protect against acute respiratory infections. Aerosols are defined as a suspension of very small (0.01–100 μm in diameter) particles or droplets in the air.[19] Studies with manikins or adult volunteers exposed to an aerosol simulating what might occur in a health care setting were considered.

Study designs assessed the use of National Institute for Occupational Safety and Health certified N95 respirators compared with surgical masks. Certification must have been under public health regulations (42 CFR part 84). Respirators certified under the former regulations (at 30 CFR part 11) were ineligible because they are no longer in use.[20] We also included data on European standard filtering facepiece (FFP2) respirators (standards EN149:2001 and EN149:2001+A1:2009) as data on N95 filtering facepiece respirators. We did not include data on elastomeric facepiece respirators because they are not in widespread use in health care settings. The term "surgical mask" was considered equivalent to medical masks, procedural masks, isolation masks, laser masks, fluid-resistant masks and face masks that meet bacterial and particle filtration efficiency standards required by the US Food and Drug Administration (ASTM standard F2100–11) but are not certifiable as N95 respirators. Other types of respirators and surgical masks not explicitly described here were excluded.

Data extraction and quality assessment

The primary outcome of interest from RCTs, cohort studies and case–control studies was laboratory-confirmed respiratory infection, including respiratory infections diagnosed by means of polymerase chain reaction, serology, respiratory virus culture and *Bordetella pertussis* bacterial culture. Secondary outcomes were influenza-like illness, and workplace absenteeism due to hospital-acquired respiratory infections. The outcomes extracted from surrogate exposure studies were filter penetration, face-seal leakage and total inward leakage.

Two reviewers (J.D.S. and C.C.M.) independently screened abstracts, titles and full texts as described in the selection of studies. Data extraction was conducted using an electronic spreadsheet template (completed independently by J.D.S. and C.C.M.) and compared for discrepancies. Data from surrogate exposure studies were transformed, when appropriate, from fit-factors, protection factors or filter efficiencies to penetration percentages. When necessary, one of us (J.D.S.) contacted authors for additional information (Appendix 1, Table S11).

Randomized controlled trials were explicitly assessed for bias according to the Cochrane risk-of-bias tool.[21] Cohort and case–control studies were assessed for risk of design-specific bias using the relevant Newcastle–Ottawa Scale.[22]

Outcome-specific quality of the body of evidence was assessed in duplicate by the same 2 reviewers using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework.[23,24] Disagreements were resolved through consultation with a third reviewer

(J.J.). The quality of evidence can be graded as high, moderate, low or very low.

Data synthesis

Where data could be combined for meta-analyses, these data were reported as odds ratios (ORs). We combined similar study designs only for the meta-analysis. Data were measured on dichotomous outcomes (laboratory-confirmed respiratory infection, influenza-like illness and workplace absenteeism). A random-effects analysis model and inverse variance statistical method were used for meta-analysis using Review Manager (RevMan).[25]

Cluster RCTs were adjusted for the meta-analysis with individual RCTs. We used the intraclass correlation coefficient to determine the design effect.[26] Design effect was used to determine the effective sample size.[26] When the effective sample size was not a whole number, it was rounded to the nearest whole number.

For meta-analyses involving rare events, zero cell counts were adjusted by including a correction (the reciprocal of the size of the contrasting study arm).[27]

We assessed evidence of heterogeneity using the χ^2 test and I^2 statistic; a χ^2 value less than 0.10 or an I^2 value greater than 50% indicated significant heterogeneity.[28,29] Subgroup analysis was planned if there were more than 5 pooled studies and when significant heterogeneity was present.

All statistical analyses were performed with the use of RevMan (version 5.2; The Nordic Cochrane Centre, The Cochrane Collaboration, 2012).

Results

Search results and study characteristics

We screened 8962 titles, excluded 8855 and retrieved 107 articles for full-text assessment. We selected 31 eligible articles involving 29 studies; 6 were clinical studies that we included in the meta-analysis, and 23 were surrogate exposure studies (Figure 1). No unpublished abstracts of RCTs, cohort studies or case–control studies were found.

We included 3 RCTs, 1 cohort study and 2 case–control studies in the meta-analysis.[11–17] The main characteristics of these studies are found in Table 1. All 6 studies reported laboratory-confirmed respiratory infection. Definitions of laboratory-confirmed respiratory infection differed. None of the RCTs used *B. pertussis* bacterial culture or viral culture. Neither of the RCTs by MacIntyre and colleagues[12–14] used serology. The SARS cases in the cohort study[15] and one of the case–control studies were confirmed only by serology.[16] Zhang and colleagues[17] confirmed influenza only by polymerase chain reaction. All of the RCTs reported on influenza-like illness. One RCT also reported workplace absenteeism; however, the outcome could not be confirmed to result from nosocomial respiratory infections.[11]

Effect on outcomes

No significant difference in risk of laboratory-confirmed respiratory infection was detected between health care workers using N95 respirators and those using surgical masks in the meta-analysis of the RCTs (OR 0.89, 95% confidence interval [CI] 0.64–1.24; $I^2 = 0\%$), the cohort study (OR 0.43, 95% CI 0.03–6.41) or the case–control studies (OR 0.91, 95% CI 0.25–3.36; $I^2 = 0\%$) (Figure 2). Similar results were found in 2 post-hoc meta-analyses: in one, we combined data from the 3 observational studies (OR 0.79, 95% CI 0.24–2.56; $I^2 = 0\%$); in the other, although not advised, we pooled data from all of the studies as an intellectual exercise to try to ascertain whether

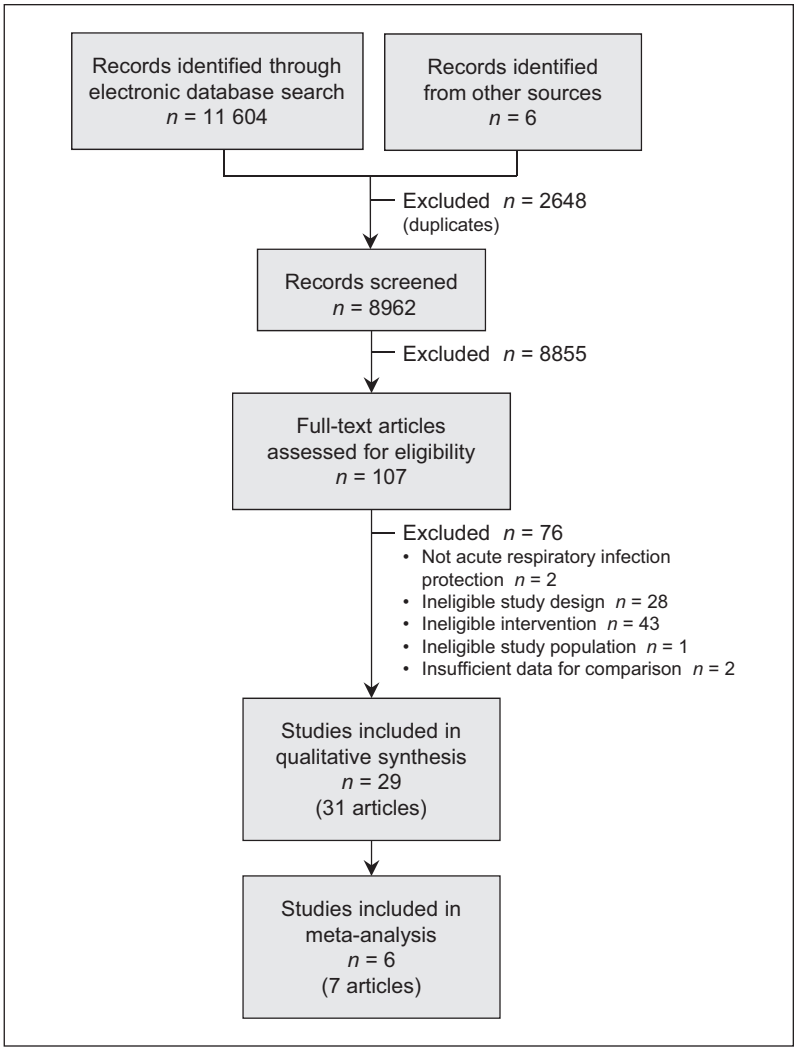


Figure 1: Selection of studies for the meta-analysis.

more precision could theoretically be obtained (OR 0.88, 95% CI 0.64–1.21; $I^2 = 0\%$).

We found no significant difference in risk of influenza-like illness between N95 respirators and surgical masks in the meta-analysis of the 3 RCTs (OR 0.51, 95% CI 0.19–1.41; $I^2 = 18\%$) (Figure 2). We also found no significant difference in risk of workplace absenteeism between N95 respirators and surgical masks in the 1 RCT that measured this outcome[11] (OR 0.92, 95% CI 0.57–1.50) (Figure 2).

Risk of bias

The risk of bias for the RCTs is summarized in Figure S1 of Appendix 1. In brief, risk-of-bias ratings were identical across each domain of the Cochrane risk-of-bias tool for all included RCTs (low risk of bias for random sequence

Table 1: Characteristics of studies included in the meta-analysis[11–17]

Study	Setting	Participants	Outcomes	Interventions	Notes
Randomized controlled trials					
Loeb et al., 2009[11]	8 hospitals in Ontario, Canada: emergency departments, acute medical units and pediatric units	446 nurses; individual-level randomization	• Laboratory-confirmed respiratory infection, influenza-like illness, workplace absenteeism • 5-wk follow-up	• Intervention: targeted use, fit-tested N95 respirator • Control: targeted use, surgical mask	• Noninferiority trial • Detection of influenza A and B, respiratory syncytial virus metapneumovirus, parainfluenza virus, rhinovirus–enterovirus, coronavirus and adenovirus
MacIntyre et al., 2011/2014[12,13]	15 hospitals in Beijing: emergency departments and respiratory wards	1441 nurses, doctors and ward clerks; cluster randomization by hospital	• Laboratory-confirmed respiratory infection, influenza-like illness • 5-wk follow-up	• Intervention 1: continual use, fit-tested N95 respirator • Intervention 2: continual use, non–fit-tested N95 respirator • Control: continual use, surgical mask	Detection of influenza A and B, respiratory syncytial virus metapneumovirus, parainfluenza virus, rhinovirus–enterovirus, coronavirus, adenovirus, *Streptococcus pneumoniae*, *Bordetella pertussis*, *Chlamydophila pneumoniae*, *Mycoplasma pneumoniae* and *Haemophilus influenzae* type B
MacIntyre et al., 2013[14]	19 hospitals in Beijing: emergency departments and respiratory wards	1669 nurses, doctors and ward clerks; cluster randomization by ward	• Laboratory-confirmed respiratory infection, influenza-like illness • 5-wk follow-up	• Intervention 1: continual use, fit-tested N95 respirator • Intervention 2: targeted use, fit-tested N95 respirator • Control: continual use, surgical mask	Detection of influenza A and B, respiratory syncytial virus metapneumovirus, parainfluenza virus, rhinovirus–enterovirus, coronavirus, adenovirus, *S. pneumoniae*, *B. pertussis*, *C. pneumoniae*, *M. pneumoniae* and *H. influenzae* type B
Cohort study					
Loeb et al., 2004[15]	2 hospitals in Ontario: coronary care units and ICUs with SARS patients	43 nurses	Laboratory-confirmed respiratory infection	• Intervention: N95 respirator • Control: surgical mask	• Retrospective • Only 20 nurses reported exposures and consistent use of facial protective equipment • Detection of SARS
Case–control studies					
Seto et al., 2003[16]	5 hospitals in Hong Kong: emergency departments and medicine units	13 infected (cases) and 241 noninfected (controls) nurses, doctors, health care assistants and domestic staff	Laboratory-confirmed respiratory infection	• N95 respirator • Surgical mask • Paper mask	• No cases in N95 respirator or surgical mask groups • 143 controls wore either surgical mask or N95 respirator • Detection of SARS
Zhang et al., 2013[17]	25 hospitals in Beijing: emergency departments, respiratory wards, ICUs, outpatient departments, technical clinic departments and management	51 infected (cases) and 204 noninfected (controls) doctors, nurses, technicians and other	Laboratory-confirmed respiratory infection	• N95 respirator • Surgical mask • Cloth mask	• Cases and controls matched 1:4 by hospital, ward, age and sex • 40 cases wore either N95 respirator or surgical mask • 159 controls wore either surgical mask or N95 respirator • Detection of pandemic H1N1 influenza virus

Note: ICU = intensive care unit, SARS = severe acute respiratory syndrome.

generation, incomplete outcome data, selective reporting and "other" bias; unclear risk of bias for allocation concealment; and high risk of bias for blinding of participants) except for blinding of outcome assessment, which was rated as unclear risk of bias for the RCT by

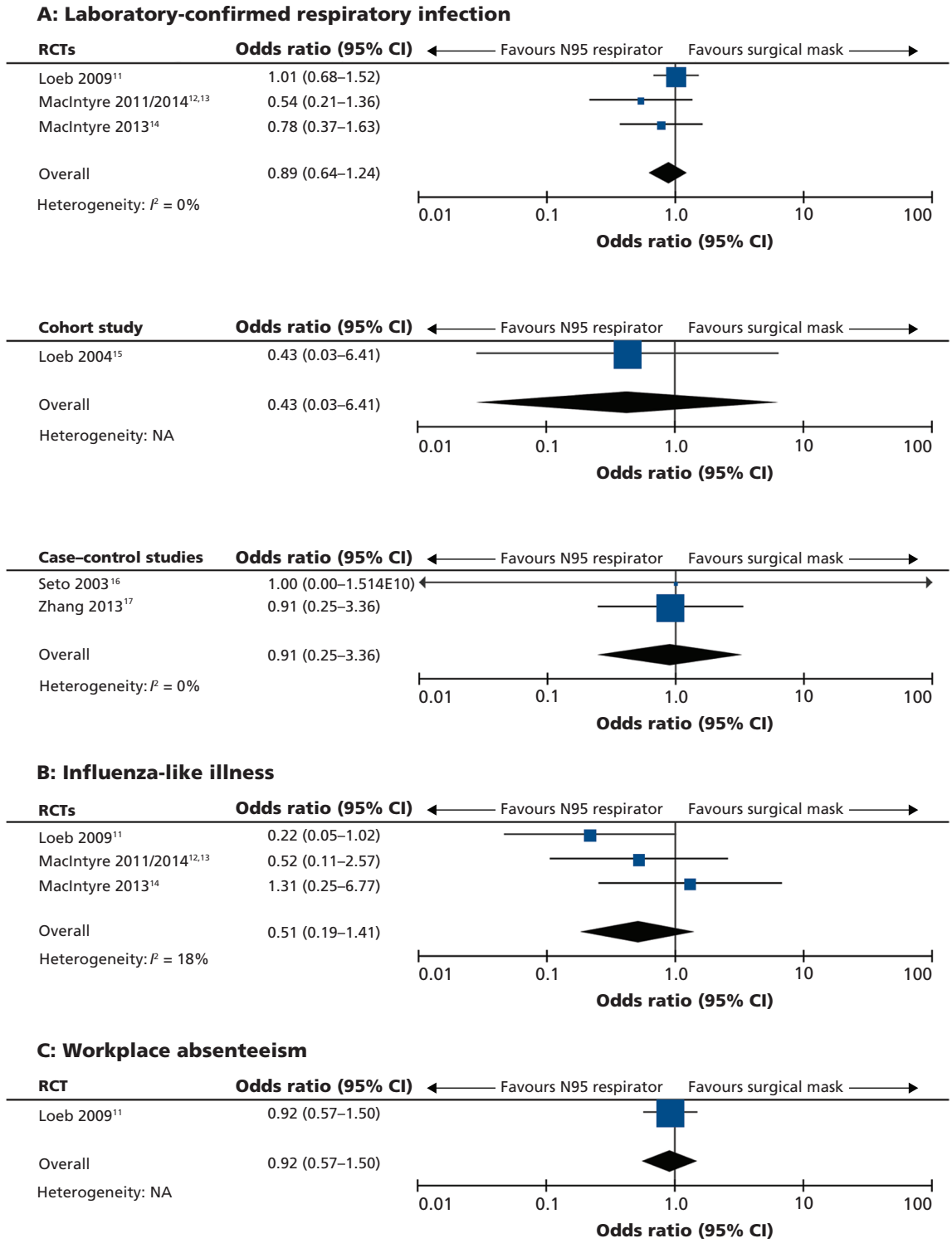


Figure 2: Results of meta-analysis to determine effectiveness of N95 respirators versus surgical masks in protecting health care workers against acute respiratory infection. Outcomes were (A) laboratory-confirmed respiratory infection, (B) influenza-like illness and (C) workplace absenteeism. Values less than 1.0 favour N95 respirator. CI = confidence interval, NA = not applicable, RCT = randomized controlled trial.

Loeb and colleagues[11] but as high risk of bias for the other 2 RCTs.[12–14]

Risk of bias for the cohort and case–control studies is summarized in Table S12 of Appendix 1. In brief, the cohort study[15] received a rating of 6 stars, one of the case–control studies received 3 stars,[16] and the other case–control study received 6 stars.[17]

Outcome-specific quality of evidence

The ratings of importance and outcome-specific quality of evidence that we assessed using the GRADE approach are summarized in Table S13 of Appendix 1. In brief, laboratory-confirmed respiratory infection was deemed a critically important outcome for decision-making with low-quality evidence from RCTs, and an important outcome for decision-making with very-low-quality evidence from observational studies. Influenza-like illness was rated as an important outcome for decision-making with very-low-quality evidence from RCTs. Work-related absenteeism was considered not an important outcome for decision-making with very-low-quality evidence from 1 RCT.

We did not conduct subgroup analyses because no significant heterogeneity was detected. No meaningful sensitivity analyses could be performed because too few studies were included.

Summary of surrogate exposure studies

Twenty-three surrogate exposure studies were included.[30–53] Their outcomes and general methods (e.g., participants, particles used for exposure, number and type of respirator or surgical mask used, flow rates and breathing rates of manikins, size of challenge particles and range of particle size measured) are summarized in Appendix 2 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.150835/-/DC1). In general, compared with surgical masks, N95 respirators showed less filter penetration, less face-seal leakage and less total inward leakage under the laboratory experimental conditions described.

Interpretation

Results of our systematic review and meta-analysis show that there was no significant difference between N95 respirators and surgical masks when used by health care workers to prevent transmission of acute respiratory infections from patients. However, wide 95% CIs from our meta-analysis must be interpreted as insufficient evidence to determine whether there is a clinically significant difference. Findings from the surrogate exposure studies suggest that N95 respirators are superior to surgical masks for filter penetration, face-seal leakage and total inward leakage under laboratory conditions.

It was not surprising to find that N95 respirators were generally more efficient filters with better face-seal characteristics than surgical masks when tested in the laboratory. However, transmission of acute respiratory infections is a complex process that may not be appropriately replicated by surrogate exposure studies. Because the face seal is important for the efficiency of the N95 respirator, fit-testing is recommended for health care workers.[2] N95 respirators are often considered uncomfortable for regular use, and improper wearing or adjustment of the respirator because of discomfort could lead to inadvertent face contamination, thus negating the potential protective benefit.[54,55] Furthermore, we do not have an adequate understanding of the number, size and dispersion of the droplets that contain live, infectious particles produced by infected patients.[56] A laboratory-based study reported data that humans infected with influenza rarely produce aerosols that contain infectious viral particles.[57] In 2 other laboratory studies, participants infected with influenza produced droplets containing viral RNA, but viral RNA could not be detected on manikin headforms or on filters of breathing manikins at distances as close as 0.1 m following participants breathing, counting, coughing or laughing.[7]

Limitations

Despite our study's many strengths, including a comprehensive search strategy for published data and grey literature, and a thorough review and assessment for risk of bias and quality of evidence using validated tools, limitations of this review should be acknowledged.

None of the studies included in the meta-analysis, except the RCT by Loeb and colleagues,[11] independently audited compliance with the intervention. Potential confounding due to concurrent interventions (e.g., gloves, gowns and hand hygiene practices) as part of routine and additional precautions for droplet transmission were not accounted for by our meta-analysis.

We did not assess the impact of harms associated with mask and respirator use that could negatively affect the efficacy of the assigned intervention because it was out of the scope of our review.[55]

Acute respiratory infections may have been acquired during the study from community exposures rather than nosocomial exposure. In one of the RCTs,[12,13] transmission may have occurred via contamination of provided respiratory protective equipment during storage and reuse of masks and respirators throughout the workday.

Only 2 respiratory virus seasons were assessed by the 3 RCTs; in one trial,[14] the peak period of one of these influenza seasons was missed, and in another trial,[11] the H1N1 outbreak in 2009 halted the study during the other respiratory season. Year-to-year strain variation of influenza necessitates additional data from other seasons during peak periods.

The weighting of the meta-analysis was influenced by the laboratory-confirmed respiratory infection outcome of serology used in one of the RCTs.[11] However, health care workers who received influenza vaccination were appropriately excluded from analysis based only on serology.

Bias due to lack of blinding in all studies was a key factor in the relatively low GRADE quality assessment, and it is impossible to overcome because the health care workers would know which mask they were wearing.

Finally, these results are not generalizable to infections transmitted primarily through airborne routes (i.e., tuberculosis, measles and varicella) or to protection from acute respiratory infections during aerosol-generating medical procedures.[3]

Conclusion

Although N95 respirators appeared to have a protective advantage over surgical masks in laboratory settings, our meta-analysis showed that there were insufficient data to determine definitively whether N95 respirators are superior to surgical masks in protecting health care workers against transmissible acute respiratory infections in clinical settings. Additional, large RCTs are needed to detect a potentially clinically important difference owing to small event rates. Initial guidelines on preventing acute respiratory infection relied on surrogate exposure data and data extrapolated from the protection of health care workers against tuberculosis because clinical evidence did not exist at that time.[58,59] Randomized controlled trials conducted in clinical settings represent the most valid information to evaluate the effectiveness of N95 respirators. They are more relevant to real clinical situations and report actual outcomes in health care workers, and therefore they are the best evidence on effectiveness to inform policy-making.

References

1. Siegel JD, Rhinehart E, Jackson M, et al. 2007 guideline for isolation precautions: preventing transmission of infectious agents in health care settings. *Am J Infect Control* 2007;35:S65-164.
2. *Routine practices and additional precautions: in all health care settings, 3rd ed.* In: Ontario Agency for Health Protection and Promotion, Provincial Infectious Diseases Advisory Committee. Toronto: Queen's Printer for Ontario; 2012. Available: www.publichealthontario.ca/en/eRepository/RPAP_All_HealthCare_Settings_Eng2012.pdf (accessed 2014 Dec. 12).
3. Infection prevention and control of epidemic- and pandemic-prone acute respiratory infections in health care. Geneva: World Health Organization; 2014. Available: www.who.int/csr/bioriskreduction/infection_control/publication/en/ (accessed 2014 Dec. 9).
4. Chung SJ, Ling ML, Seto WH, et al. Debate on MERS-CoV respiratory precautions: Surgical mask or N95 respirators? *Singapore Med J* 2014;55:294-7.
5. Brankston G, Gitterman L, Hirji Z, et al. Transmission of influenza A in human beings. *Lancet Infect Dis* 2007;7:257-65.
6. Hall CB. The spread of influenza and other respiratory viruses: complexities and conjectures. *Clin Infect Dis* 2007;45:353-9.
7. Tang JW, Gao CX, Cowling BJ, et al. Absence of detectable influenza RNA transmitted via aerosol during various human respiratory activities — experiments from Singapore and Hong Kong. *PLoS One* 2014;9:e107338.
8. Tellier R. Review of aerosol transmission of influenza A virus. *Emerg Infect Dis* 2006;12:1657-62.
9. Roy CJ, Milton DK. Airborne transmission of communicable infection — the elusive pathway. *N Engl J Med* 2004;350:1710-2.
10. Clynes N. Surgical masks vs N95 respirators for preventing influenza. *JAMA* 2010;303:937-8, author reply 938-9.
11. Loeb M, Dafoe N, Mahony J, et al. Surgical mask vs N95 respirator for preventing influenza among health care workers: a randomized trial. *JAMA* 2009;302:1865-71.
12. MacIntyre CR, Wang Q, Cauchemez S, et al. A cluster randomized clinical trial comparing fit-tested and non-fit-tested N95 respirators to medical masks to prevent respiratory virus infection in health care workers. *Influenza Other Respir Viruses* 2011;5:170-9.
13. MacIntyre CR, Wang Q, Rahman B, et al. Efficacy of face masks and respirators in preventing upper respiratory tract bacterial colonization and co-infection in hospital healthcare workers. *Prev Med* 2014;62:1-7.
14. MacIntyre CR, Wang Q, Seale H, et al. A randomized clinical trial of three options for N95 respirators and medical masks in health workers. *Am J Respir Crit Care Med* 2013;187:960-6.
15. Loeb M, McGeer A, Henry B, et al. SARS among critical care nurses, Toronto. *Emerg Infect Dis* 2004;10:251-5.
16. Seto WH, Tsang D, Yung RW, et al. Effectiveness of precautions against droplets and contact in prevention of nosocomial transmission of severe acute respiratory syndrome (SARS). *Lancet* 2003;361:1519-20.
17. Zhang Y, Seale H, Yang P, et al. Factors associated with the transmission of pandemic (H1N1) 2009 among hospital healthcare workers in Beijing, China. *Influenza Other Respir Viruses* 2013;7:466-71.
18. Chughtai AA, Seale H, MacIntyre CR. Availability, consistency and evidence-base of policies and guidelines on the use of mask and respirator to protect hospital health care workers: a global analysis. *BMC Res Notes* 2013;6:216.
19. National Institute for Occupational Safety and Health (NIOSH). *Workplace safety and health topics: aerosols*. Atlanta: US Centers for Disease Control and Prevention; 2010. Available: www.cdc.gov/niosh/topics/aerosols/ (accessed 2014 Dec. 10).
20. National Institute for Occupational Safety and Health (NIOSH). *NIOSH guide to the selection and use of particulate respirators*. Atlanta: US Centers for Disease Control and Prevention; 1996. Available: www.cdc.gov/niosh/docs/96-101/ (accessed 2015 Oct. 20).
21. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
22. Wells GA, Shea B, O'Connell D, et al. *The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. Ottawa: The Ottawa Hospital; 2011. Available: www.ohri.ca/programs/clinical_epidemiology/oxford.asp (accessed 2014 Oct. 22).
23. GRADEpro [computer program]. Hamilton (ON): McMaster University; 2014. Available: www.guidelinedevelopment.org (accessed 2014 Nov. 20).
24. GRADE Working Group. Grading the quality of evidence and the strength of recommendations. Available: www.gradeworkinggroup.org/intro.htm (accessed 2014 Apr. 30).
25. Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions, version 5.1.0.* Oxford (UK): Cochrane Collaboration; 2011. Available: handbook.cochrane.org/ (accessed 2014 Jan. 26).
26. Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann Fam Med* 2004;2:204-8.
27. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med* 2004;23:1351-75.
28. Higgins JP, Thomphson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539-58.
29. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-60.
30. Bałazy A, Toivola M, Adhikari A, et al. Do N95 respirators provide 95% protection level against airborne viruses, and how adequate are surgical masks? *Am J Infect Control* 2006;34:51-7.

31. Bischoff WE, Reid T, Russell GB, et al. Transocular entry of seasonal influenza-attenuated virus aerosols and the efficacy of N95 respirators, surgical masks, and eye protection in humans. *J Infect Dis* 2011;204:193-9.
32. Checchi L, Montevecchi M, Moreschi A, et al. Efficacy of three face masks in preventing inhalation of airborne contaminants in dental practice. *J Am Dent Assoc* 2005;136:877-82.
33. Davidson CS, Green CF, Gibbs SG, et al. Performance evaluation of selected N95 respirators and surgical masks when challenged with aerosolized endospores and inert particles. *J Occup Environ Hyg* 2013;10:461-7.
34. Davidson C, Green CF, Panlilio AL, et al. Method for evaluating the relative efficiency of selected N95 respirators and surgical masks to prevent the inhalation of airborne vegetative cells by healthcare personnel. *Indoor Built Environ* 2011;20:265-77.
35. Derrick JL, Li PT, Tang SP, et al. Protecting staff against airborne viral particles: in vivo efficiency of laser masks. *J Hosp Infect* 2006;64:278-81.
36. Diaz KT, Smaldone GC. Quantifying exposure risk: surgical masks and respirators. *Am J Infect Control* 2010;38:501-8.
37. Duling MG, Lawrence RB, Slaven JE, et al. Simulated workplace protection factors for half-facepiece respiratory protective devices. *J Occup Environ Hyg* 2007;4:420-31.
38. Gawn J, Clayton M, Makison C, et al. *Evaluating the protection afforded by surgical masks against influenza bioaerosols: gross protection of surgical masks compared to filtering facepiece respirators*. Norwich (UK): Health and Safety Executive; 2008. Available: www.hse.gov.uk/research/rrpdf/rr619.pdf (accessed 2014 June 3).
39. Grinshpun SA, Haruta H, Eninger RM, et al. Performance of an N95 filtering facepiece particulate respirator and a surgical mask during human breathing: two pathways for particle penetration. *J Occup Environ Hyg* 2009;6:593-603.
40. He X, Reponen T, McKay RT, et al. Effect of particle size on the performance of an N95 filtering facepiece respirator and a surgical mask at various breathing conditions. *Aerosol Sci Technol* 2013;47:1180-7.
41. He X, Reponen T, McKay R, et al. How does breathing frequency affect the performance of an N95 filtering facepiece respirator and a surgical mask against surrogates of viral particles? *J Occup Environ Hyg* 2014;11:178-85.
42. Lawrence RB, Duling MG, Calvert CA, et al. Comparison of performance of three different types of respiratory protection devices. *J Occup Environ Hyg* 2006;3:465-74.
43. Lee SA, Grinshpun SA, Reponen T. Respiratory performance offered by N95 respirators and surgical masks: human subject evaluation with NaCl aerosol representing bacterial and viral particle size range. *Ann Occup Hyg* 2008;52:177-85.
44. Li Y, Wong T, Chung J, et al. In vivo protective performance of N95 respirator and surgical facemask. *Am J Ind Med* 2006;49:1056-65.
45. Lindsley WG, King WP, Thewlis RE, et al. Dispersion and exposure to a cough-generated aerosol in a simulated medical examination room. *J Occup Environ Hyg* 2012;9:681-90.
46. Mansour MM, Smaldone GC. Respiratory source control versus receiver protection: impact of facemask fit. *J Aerosol Med Pulm Drug Deliv* 2013;26:131-7.
47. Mitakakis TZ, Tovey ER, Yates DH, et al. Particulate masks and non-powdered gloves reduce latex allergen inhaled by healthcare workers. *Clin Exp Allergy* 2002;32:1166-9.
48. Noti JD, Lindsley WG, Blachere FM, et al. Detection of infectious influenza virus in cough aerosols generated in a simulated patient examination room. *Clin Infect Dis* 2012;54:1569-77.
49. Qian Y, Willeke K, Grinshpun SA, et al. Performance of N95 respirators: filtration efficiency for airborne microbial and inert particles. *Am Ind Hyg Assoc J* 1998;59:128-32.
50. Rengasamy S, Eimer BC. N95-companion measurement of C_{out}/C_{in} ratios for two N95 filtering facepiece respirators and one surgical mask. *J Occup Environ Hyg* 2013;10:527-32.
51. Rengasamy S, Eimer BC, Szalajda J. A quantitative assessment of the total inward leakage of NaCl aerosol representing submicron size bioaerosol through N95 filtering facepiece respirators and surgical masks. *J Occup Environ Hyg* 2014;11:388-96.
52. Wen Z, Yu L, Yang W, et al. Assessment the protection performance of different level personal respiratory protection masks against viral aerosol. *Aerobiologia* 2013;29:365-72.
53. Zou Z, Yao M. Airflow resistance and bio-filtering performance of carbon nanotube filters and current facepiece respirators. *J Aerosol Sci* 2015;79:61-71.
54. Rebmann T, Carrico R, Wang J. Physiologic and other effects and compliance with long-term respirator use among medical intensive care unit nurses. *Am J Infect Control* 2013;41:1218-23.
55. Jefferson T, Del Mar CB, Dooley L, et al. Physical interventions to interrupt or reduce the spread of respiratory viruses. *Cochrane Database Syst Rev* 2011;(7):CD006207.
56. Morawska L. Droplet fate in indoor environments, or can we prevent the spread of infection? *Indoor Air* 2006;16:335-47.
57. Milton DK, Fabian MP, Cowling BJ, et al. Influenza virus aerosols in human exhaled breath: particle size, culturability, and effect of surgical masks. *PLoS Pathog* 2013;9:e1003205.
58. Bozzi CJ, Burwen DR, Dooley SW, et al. Guidelines for preventing the transmission of *Mycobacterium tuberculosis* in health-care facilities, 1994. *MMWR Recomm Rep* 1994;43(RR-13):1-132.
59. Gammaitoni L, Nucci MC. Using a mathematical model to evaluate the efficacy of TB control measures. *Emerg Infect Dis* 1997;3:335-42.

Affiliations: Public Health Ontario (Smith, MacDougall, Johnstone, Copes, Schwartz, Garber); Department of Medicine (Johnstone), University of Toronto; Dalla Lana School of Public Health (Copes, Schwartz), University of Toronto, Toronto, Ont.; Department of Medicine (Garber), University of Ottawa, Ottawa, Ont.

Contributors: Jeffrey Smith drafted the protocol and data extraction template, and performed the literature screening, data extraction and analysis. Colin MacDougall provided input for all stages of the project, and performed screening and data extraction. Jennie Johnstone provided input and made revisions on the protocol, data extraction template and analyses. Ray Copes, Brian Schwartz and Gary Garber were involved in the study conception. Gary Garber oversaw the project. Jeffrey Smith drafted the manuscript, and Colin MacDougall, Jennie Johnstone, Ray Copes, Brian Schwartz and Gary Garber provided critical feedback on the manuscript. All of the authors approved the final version to be published and agreed to act as guarantors of the work.

Acknowledgement: The authors are grateful for the contributions of the Library Services team at Public Health Ontario in the development, refinement and execution of the search strategy for this review.

Accepted: 22 July 2017

DOI: 10.1111/irv.12474

ORIGINAL ARTICLE

WILEY

The efficacy of medical masks and respirators against respiratory infection in healthcare workers

Chandini Raina MacIntyre[1,2] | Abrar Ahmad Chughtai[1] | Bayzidur Rahman[1] | Yang Peng[3] | Yi Zhang[3] | Holly Seale[1] | Xiaoli Wang[3] | Quanyi Wang[3]

[1]School of Public Health and Community Medicine, University of New South Wales, Sydney, NSW, Australia

[2]College of Public Service & Community Solutions, and College of Health Solutions, Arizona State University, Phoenix, AZ, USA

[3]The Beijing Centre for Disease Prevention and Control, Beijing, China

Correspondence
Abrar Ahmad Chughtai, School of Public Health and Community Medicine, UNSW Medicine, University of New South Wales, Sydney, NSW, Australia.
Email: abrar.chughtai@unsw.edu.au

Funding information
Trial 1: 3M helped in fit testing, no financial support was provided.
Trial 2: Australian National Health & Medical Research Council of Australia (Grant # 630787).
Funders were not involved in the design and implementation of the study

Objective: We aimed to examine the efficacy of medical masks and respirators in protecting against respiratory infections using pooled data from two homogenous randomised control clinical trials (RCTs).

Methods: The data collected on 3591 subjects in two similar RCTs conducted in Beijing, China, which examined the same infection outcomes, were pooled. Four interventions were compared: (i) continuous N95 respirator use, (ii) targeted N95 respirator use, (iii) medical mask use and (iv) control arm. The outcomes were laboratory-confirmed viral respiratory infection, influenza A or B, laboratory-confirmed bacterial colonisation and pathogens grouped by mode of transmission.

Results: Rates of all outcomes were consistently lower in the continuous N95 and/or targeted N95 arms. In adjusted analysis, rates of laboratory-confirmed bacterial colonisation (RR 0.33, 95% CI 0.21-0.51), laboratory-confirmed viral infections (RR 0.46, 95% CI 0.23-0.91) and droplet-transmitted infections (RR 0.26, 95% CI 0.16-0.42) were significantly lower in the continuous N95 arm. Laboratory-confirmed influenza was also lowest in the continuous N95 arm (RR 0.34, 95% CI 0.10-1.11), but the difference was not statistically significant. Rates of laboratory-confirmed bacterial colonisation (RR 0.54, 95% CI 0.33-0.87) and droplet-transmitted infections (RR 0.43, 95% CI 0.25-0.72) were also lower in the targeted N95 arm, but not in medical mask arm.

Conclusion: The results suggest that the classification of infections into droplet versus airborne transmission is an oversimplification. Most guidelines recommend masks for infections spread by droplets. N95 respirators, as "airborne precautions," provide superior protection for droplet-transmitted infections. To ensure the occupational health and safety of healthcare worker, the superiority of respirators in preventing respiratory infections should be reflected in infection control guidelines.

KEYWORDS
droplet infections, healthcare workers, influenza, masks, medical masks, respirators

1 | BACKGROUND

There is currently a lack of consensus around the efficacy of medical masks and respirators for healthcare workers (HCWs) against influenza, with only five published randomised control trials (RCTs) in HCWs conducted to date.[1-5] While N95 respirators have been shown to be superior to medical masks in preventing clinical respiratory infection (CRI), influenza illness (ILI) and other outcomes, none of the

studies were adequately powered to examine laboratory-confirmed influenza.

In the smallest of the trials, involving only 32 HCWs, there was no difference in the rates of respiratory illnesses between HCWs who used medical masks and the control group.[1] A Canadian study of 422 hospital nurses compared targeted use of N95 respirators and medical masks and found that the rate of serologically defined influenza was 25% in both arms.[2] However, in the absence of a control arm for comparison, the finding of no difference in outcomes between the intervention arms could represent either equal efficacy or equal inefficacy of the two interventions. The other two published HCW RCTs used a more specific and less sensitive definition of influenza based on nucleic acid testing (NAT) of respiratory specimens in symptomatic subjects. As such, even these substantially larger RCTs were unable to demonstrate any significant difference in influenza infection between N95 respirators and medical masks.[3,4] Finally, a recent study examined the efficacy of cloth masks compared to medical mask and control groups, and found that cloth masks may increase the risk of infection in HCWs.[5]

Guidelines for respiratory protection have been driven by presumed transmission mode alone, and under an assumption that influenza and other pathogens are spread by one mode alone.[6] However, the paradigm of unimodal droplet or airborne spread is based on outmoded experiments from the 1940s, which concluded that only large droplets are found at close proximity to the patient, while small droplet nuclei and airborne particles are found at a longer distance.[7-9] It has since been shown that both small and large particles can exist at short distances from the patient, and that aerosolised transmission can occur at close proximity.[9]

In our two published RCTs conducted in China,[3,4] we used the same outcomes, case definitions and measurement tools, and used the same testing methods for a range of different pathogens transmitted by different routes. This afforded an opportunity to pool the data from both trials for improved statistical power to examine the outcomes by pathogens and mode of transmission. The aim of this pooled analysis was to examine the efficacy of medical masks and respirators in HCWs against respiratory infection.

2 | METHODS

We pooled the results of our two RCTs on mask and respirator use in hospital HCWs in Beijing, China. The first RCT (Trial 1) was conducted from December 2008 to January 2009,[3] and included 1441 HCWs randomised to: medical mask arm (n = 492), N95 fit-tested arm (n = 461) and N95 non-fit-tested arm (n = 488). The rate of fit-test failure was very low (5/461) in this trial, so data from both N95 arms were combined for analysis.

An additional 481 healthcare workers from nine hospitals were recruited to a control arm. These hospitals were purposefully selected as they indicated low levels of routine mask/respirator use during a pre-trial assessment. Participants in the control arms continued their usual mask wearing practices and were followed using the same protocol as applied to the other arms.[3]

The second trial (Trial 2) was conducted from 28 December 2009 to 7 February 2010, using the same design.[4] In Trial 2, participants were randomised to three arms: medical masks at all times on shift (n = 572), continuous N95 respirators at all times on shift (n = 516) and targeted/intermittent use of N95 respirators only while doing high-risk procedures or barrier nursing of a patient with known respiratory illness (n = 521). Fit testing was not performed in the second RCT. In both trials, participants were followed for 4 weeks of wearing the medical masks or respirators, and an extra week of non-wearing of masks for the development of symptoms. Demographic and clinical data were collected, including gender, age, smoking, vaccination status, pre-existing medical illnesses, hand hygiene and high-risk procedures. Pharyngeal swabs were collected from symptomatic participants, and samples were tested at the laboratories of the Beijing Centers for Disease Control and Prevention. There was no major difference in the products used in both clinical trials. In the first trial, we used medical masks (3M, catalogue number 1820) and N95 fit/non-fit-tested respirator (3M, catalogue number 9132). The following products were used in the second trial: medical masks (3M, catalogue number 1817) and respirator (3M, catalogue number 1860).

The interventions compared in the pooled analysis were as follows: (i) continuous use of N95 respirators (pooled data from both trials - 1530 subjects); (ii) targeted N95 respirator use (data from trial 2-516 subjects); (iii) continuous use of medical masks (pooled data from both trials - 1064 subjects) and (iv) and a control group (data from trial 1-481 subjects).

Only laboratory-confirmed outcomes were included in the analysis, which were defined and measured identically in both trials, and comprised: (i) laboratory-confirmed viral respiratory infection (detection of adenoviruses; human metapneumovirus; coronavirus 229E / NL63; parainfluenza viruses 1, 2 and 3; influenza viruses A and B; respiratory syncytial virus A and B; rhinovirus A/B and coronavirus OC43 /HKU1 by multiplex PCR); (ii) laboratory-confirmed (multiplex PCR) influenza A or B and (iii) laboratory-confirmed bacterial colonisation (Streptococcus pneumonia, Haemophilus influenza, Bordetella pertussis, Chlamydophila pneumoniae and Mycoplasma pneumonia).[3,4] The laboratory testing has previously been described.[3,4]

Laboratory-confirmed bacteria and viruses identified in participants were categorised according to droplet (n = 285), contact (n = 6) and airborne (n = 3) transmission modes (Table S1A). Sixty-one co-infection cases with multitransmission were categorised separately. Among the viruses isolated, coronavirus and influenza A/B were included in the droplet category (and thus included in the additional analysis); rhinovirus A/B was included in the airborne category and adenovirus; parainfluenza virus and respiratory syncytial virus (RSV) were included in contact category in the base case analysis. All bacteria were categorised into the droplet transmission category. For consistency, data on the transmission modes were taken from the Pathogen Safety Data Sheets (PSDSs) of the Public Health Agency of Canada[10] (Table S1A). As the largest number of confirmed infections was in the droplet category, we conducted a subgroup analysis of droplet-transmitted infections. Given there were a large number of RSV cases (n = 33) in our data set and RSV is variously categorised as

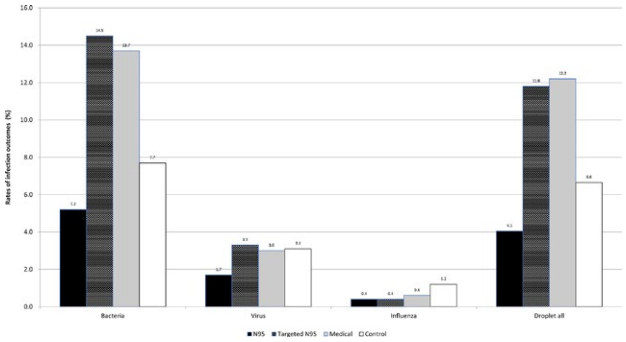


FIGURE 1 Rate of infections reported in HCWs in the different arms from Trials 1 and 2

either "droplet"[11] or "contact" spread[12] in different guidelines, we performed a sensitivity analysis by including RSV into the droplet transmission category instead of contact.

2.1 | Ethics

Ethics approvals of two clinical trials were obtained from the Institutional Review Board and Human Research Ethics Committee of the Beijing Center for Disease Prevention and Control.

2.2 | Patient involvement

We did not involve patients and their families in the design and conduct of the study. We have acknowledged the support of participants, and the results will be published in open access journal.

2.3 | Statistical analysis

The data sets from the two trials were pooled incorporating the common variables. We calculated the attack rate (proportion of outcome) of each of the four outcomes by the study arms.

We conducted a fixed effect individual patient data (IPD) meta-analysis by fitting a multivariable log binomial model, using generalised estimating equation (GEE) to account for clustering by hospital/ward. We used a fitted fixed effect model because there are only two trials. Two studies were conducted in the same setting with similar participant characteristics, and they examined the same underlying effect. In the analysis, relative risk (RR) was estimated using the control arm as the referent category after adjusting for potential confounders and their interaction terms with a trial ID number. The overall rates of seasonal infection were higher in the second trial than the first. The consistency assumption (ie between study homogeneity) for the IPD meta-analysis was tested by fitting an interaction term between trial ID and trial arms where a significant interaction is indicative of inconsistency.[13] Any interaction term (between trial ID and covariates other than trial arm) that was not a confounder was subsequently excluded from the model using backward elimination approach. This approach is described in detailed elsewhere.[4] We repeated the above-described methods for each of the outcomes.

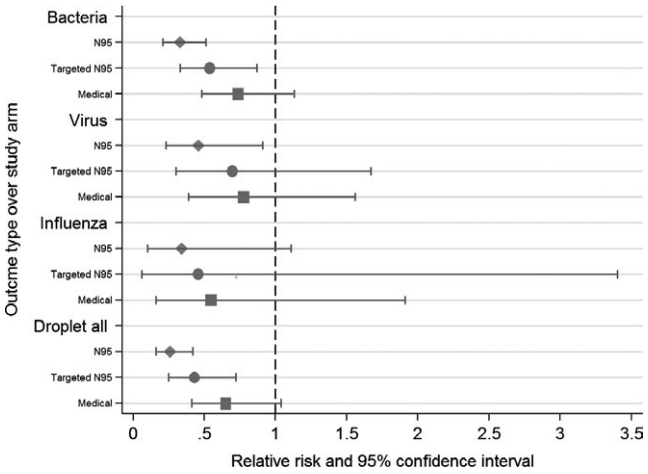


FIGURE 2 Forest plot of outcomes according to various interventions

3 | RESULTS

After combining the data sets from the two trials, 3591 cases were entered into the pooled analysis (1064 cases in the medical mask arm, 516 cases in the targeted N95 arm, 1530 cases in the continuous N95 arm and 481 cases in the control arm). The infection outcomes are presented in Figure 1. The rates of laboratory-confirmed viral respiratory infection (26/1530, 1.7%), laboratory-confirmed bacterial colonisation (79/1530, 5.2%) and droplet-transmitted infections (62/1530, 4.1%) were lowest among the continuous N95 arm. Laboratory-confirmed influenza A and B was lowest in continuous N95 (6/1530, 0.4%) and targeted N95 arms (2/516, 0.4%).

In the IPD meta-analysis, none of the interaction terms between trial arm and trial ID was significant for any of the outcome variables. Thus, the consistency assumption for the IPD meta-analysis was satisfied. However, a significant interaction was observed between trial ID and hand washing for laboratory-confirmed bacterial colonisation only; therefore, we estimated the RR for trial ID stratified by hand washing.

Figure 2 shows the forest plot of outcomes according to various interventions. All outcomes were consistently lower in the continuous N95 and targeted N95 arms. The IPD meta-analysis shows that the risk of laboratory-confirmed bacterial colonisation was lower in the

TABLE 1 Multivariable cluster adjusted log binomial model of laboratory-confirmed bacterial colonisation

Variables in the model	Relative risk (95% CI)	*P*-value
Continuous N95 arm	**0.33 (0.21-0.51)**	**<.001**
Targeted N95 arm	**0.54 (0.33-0.87)**	**.001**
Medical mask arm	0.74 (0.48-1.13)	.161
Control arm	Ref	Ref
Sex (Male)	0.60 (0.42-0.85)	.005
Trial	2.53 (1.65-3.87)	<.001
Influenza vaccine	1.13 (0.89-1.43)	.308
Trial * Hand wash	4.49 (3.12-6.48)	<.001

Bold value indicates statistically significant results.

TABLE 2 Multivariable cluster adjusted log binomial model of laboratory-confirmed viral respiratory infection

Variables in the model	Relative risk (95% CI)	*P*-value
Continuous N95 arm	**0.46 (0.23-0.91)**	**.026**
Targeted N95 arm	0.70 (0.30-1.67)	.424
Medical mask arm	0.78 (0.39-1.56)	.484
Control arm	Ref	Ref
Sex (Male)	0.69 (0.36-1.33)	.272
Hand washing	0.78 (0.51-1.20)	.264
Influenza vaccine	0.94 (0.57-1.55)	.808
Trial	1.50 (0.89-2.54)	.131

Bold value indicates statistically significant results.

TABLE 3 Multivariable cluster adjusted log binomial model of laboratory-confirmed influenza A or B

Variables in the model	Relative risk (95% CI)	*P*-value
Continuous N95 arm	0.34 (0.10-1.11)	.074
Targeted N95 arm	0.46 (0.06-3.40)	.445
Medical mask arm	0.55 (0.16-1.91)	.350
Control arm	Ref	Ref
Sex (Male)	0.27 (0.03-2.01)	.220
Hand washing	0.70 (0.29-1.73)	.446
Influenza vaccine	0.78 (0.26-2.34)	.660
Trial	0.64 (0.19-2.18)	.477

continuous N95 arm (RR 0.33, 95% CI 0.21-0.51 or 67% efficacy) and targeted N95 arm (RR 0.54, 95% CI 0.33-0.87 or 46% efficacy) (Table 1).

Laboratory-confirmed viral respiratory infections were significantly lower in the continuous N95 arm (RR 0.46, 95% CI 0.23-0.91, or 54% efficacy). The rates of laboratory-confirmed virus were also lower in the targeted N95 arm (RR 0.70, 95% CI 0.30-1.67) and medical masks arm (RR 0.78, 95% CI 0.39-1.56); however, the difference was not statistically significant (Table 2).

Laboratory-confirmed influenza was also lowest in continuous N95 arm (RR 0.34, 95% CI 0.10-1.11) but not significant (Table 3). In the subgroup analysis of droplet-transmitted infections, compared to the control arm, the efficacy of continuous N95 respirators against droplet-transmitted infections (bacterial and viral) was 74% (RR 0.26, 95% CI 0.16-0.42) and 57% in the targeted N95 arm (RR 0.43, 95% CI 0.25-0.72) (Table 4).

Inclusion of RSV cases in the droplet-transmitted pathogen category did not change the risk ratio to a large extent. If RSV cases were also included in the droplet-transmitted pathogen category, the efficacy was 70% in the continuous N95 (RR 0.30, 95% CI 0.19-0.46) and 51% in the targeted N95 arms (RR 0.49, 95% CI 0.30-0.80). The rate of droplet only transmitting viral infections was also lower in the continuous N95 and targeted N95 arms. HCWs who used a continuous N95 and targeted respirator were 85% (RR 0.15, 95% CI 0.04-0. 59) and 84% (RR 0.12, 95% CI 0.02-0.88) less likely to acquire droplet-transmitted viral infections.

TABLE 4 Multivariable cluster adjusted log binomial model of droplet-transmitted infections

Variables in the model	Relative risk (95% CI)	*P*-value
Continuous N95 arm	**0.26 (0.16-0.42)**	**<.001**
Targeted N95 arm	**0.43 (0.25-0.72)**	**.001**
Medical mask arm	0.65 (0.41-1.04)	.074
Control arm	Ref	Ref
Sex (Male)	0.63 (0.43-0.92)	.016
Hand washing	1.27 (0.99-1.62)	.068
Influenza vaccine	1.16 (0.90-1.50)	.257
Trial	3.97 (2.83-5.59)	<.001

Bold value indicates statistically significant results.

When only the continuous N95 arm was compared against control, the risk of laboratory-confirmed influenza was significantly lower in continuous N95 arm (RR 0.23 and 95% CI 0.06-0.93, or 77% efficacy). In the similar analysis, the risk of influenza was also lower in medical mask arm compared to control; however, the difference was not statistically significant (RR 0.81 and 95% CI 0.25-2.68) arm. Table 5 compares the results of this analysis with the individual studies.

4 | DISCUSSION

We demonstrated superior clinical efficacy of continuous use of N95 respirator (also known as "airborne precautions") against infections presumed to be spread by the droplet mode, including influenza. This suggests that transmission is more complex than assumed by traditional classifications, and supports the fact that both large and small droplets are present close to the patient, and that aerosol transmission may occur for presumed "droplet" infections. Respirators are designed to provide respiratory protection through filtration and fit, and properly fitted respirators provide better protection compared to medical masks.[3,4] We could not demonstrate efficacy of medical masks against any outcome, but the non-significant trend appeared to be towards protection. Medical masks may well have efficacy,[5] but if so, the degree of efficacy was too small to detect in this study, and larger studies are needed, given the widespread use of these devices in health care.

The practical implication of this research is illustrated with influenza as a case in point. Droplet and contact are thought to be primary modes of transmission for seasonal influenza; therefore, the World Health Organisation (WHO) and the Centers for Disease Control and Prevention (CDC) guidelines recommend medical masks during routine patient care, while N95 respirators are recommended during procedures in which aerosols may be generated and during other high-risk situations.[14,15] However, there is increasing evidence of aerosol transmission of influenza during routine care as well (in the absence of aerosol generating procedures), which may warrant superior respiratory protection.[16,17] Influenza research is challenging because there is high seasonal variation in activity, and the level of circulating influenza in any given year cannot be predicted when planning RCTs. In addition,

TABLE 5 Results of individual clinical trials and pooled analysis

	Arms	RCT 1 (OR/ RR)	RCT 2 (HR/ RR)	Pooled analysis
CRI	Continuous N95	0.46 (0.19-1.11)	**0.39 (0.21-0.71)**	
	Targeted N95	-	0.70 (0.39-1.24)	
	Medical masks	0.74 (0.29-1.88)	Ref	
	Control	Ref	-	
Influenza like illness	Continuous N95	0.26 (0.06-1.11)	-	
	Targeted N95	-	-	
	Medical masks	0.49 (0.12-2.07)	-	
	Control	Ref	-	
Laboratory-confirmed viruses	Continuous N95	**0.43 (0.20-0.91)**	-	**0.46 (0.23-0.91)**
	Targeted N95	-	-	0.70 (0.30-1.67)
	Medical masks	0.84 (0.38-1.85)	-	0.78 (0.39-1.56)
	Control	Ref	-	Ref
Laboratory-confirmed influenza	Continuous N95	0.25 (0.06-1.00)	-	0.34 (0.10-1.11)
	Targeted N95	-	-	0.46 (0.06-3.40)
	Medical masks	0.81 (0.25-2.68)	-	0.55 (0.16-1.91)
	Control	Ref	-	Ref
Laboratory-confirmed bacterial colonisation	Continuous N95	**0.34 (0.21-0.56)**	**0.40 (0.21-0.73)**	**0.33 (0.21-0.51)**
	Targeted N95	-	0.70 (0.40-1.24)	**0.54 (0.33-0.87)**
	Medical masks	0.67 (0.38-1.18)	Ref	0.74 (0.48-1.13)
	Control	Ref		Ref
Droplet-transmitted infections	Continuous N95	-	-	**0.26 (0.16-0.42)**
	Targeted N95	-	-	**0.43 (0.25-0.72)**
	Medical masks	-	-	0.65 (0.41-1.04)
	Control	-	-	Ref

Bold value indicates statistically significant results.

a diagnosis of influenza requires the detection of virus from respiratory specimens, or a fourfold rise in serological titres, both of which are highly resource-intensive and depend on daily subject follow-up and on optimal timing of specimen collection. For all these reasons, the published studies to date have been unable to determine whether there is a difference in efficacy against influenza infection between medical masks and N95 respirators. This study can therefore usefully inform policies for prevention of influenza.

In the first RCT, compared to medical masks, N95 respirators were found to be protective against CRI, but not against ILI or laboratory-confirmed influenza.[3] When compared with the control arm, rates of laboratory-confirmed virus and bacterial colonisation were significantly lower in N95 arm (Table 5). In the second RCT, continuous use of N95 respirators was associated with lower rates of CRI and laboratory-confirmed bacterial colonisation compared to the medical mask use.[4] Pooled analysis of these studies improved the power to analyse other infectious outcomes by intervention and to allow analysis by mode of transmission.

An important finding of this analysis was the efficacy of N95 respirators against droplet-transmitted infections. Generally, medical masks are considered sufficient for droplet-transmitted infections such as influenza.[18] However, this study has demonstrated a clear benefit of using N95 respirators (both continuous and targeted) to protect HCWs against droplet infections and does not show significant protection of medical masks. In the light of these findings, it may be prudent to use respirators when the transmission mode of a disease is unknown or when HCWs exposed to droplet-transmitted infections with a high-case fatality rate.[6] Middle East respiratory syndrome coronavirus (MERS-CoV) and Ebola virus disease (EVD) are not airborne infections, yet the CDC recommendation of using respirators to protect HCWs recognises the uncertainty around transmission.[19,20] The CDC initially recommended medical masks for Ebola, but changed their guidelines when US HCWs became infected, amidst unrest and challenges to the prior guidelines.[6,21] In contrast, the WHO recommends medical masks for MERS-CoV and Ebola [22,23] despite having older guidelines for filoviruses which recommended respirators.[24] There is a need for a more evidence-based approach to updating guidelines and ensuring consistency between different guidelines.[25]

Our study also demonstrated that, over and above the benefit of continuous use, targeted use of N95 is associated with reduced risk of infection. Many guidelines recommend targeted use,[14,15,26] and our study supports this practice. However, better protection is achieved through continuous use of respirators. This may be because HCWs cannot always identify situations in which they are at risk, especially in busy clinical settings with a high level of movement of patients and staff in and out of wards.

This study has some limitations. Firstly, the reporting of the results included in Figure 1 is different from the IPD meta-analysis results. This is due to the uneven distribution of randomisation arms and differing seasonal attack rates between the trials. In Figure 1, these between-trial differences were not taken into account. The IPD meta-analysis takes into account of these and gives an unbiased association. Secondly, the control arm in trial 1 was not randomised; however, the risk of bias is less due to similar study setting, outcome measures and participant characteristics. Moreover, whether infection was acquired in the community or the hospital cannot be determined, but the RCT design should result in community exposure being distributed equally across all arms. Finally, we categorised pathogens according to various transmission modes, while certain viruses are transmitted via multiple routes. The pooled data were suggestive of an effect of respirators against influenza, but probably did not have enough statistical power for this outcome. The major strength of this study is the use of the same endpoints, measurements and methods in the two trials, which allowed valid pooling of the data.

5 | CONCLUSION

It is a long-held belief in hospital infection control that a mask is adequate for droplet-transmitted infections. We showed that the use of respirators provides better protection against respiratory infections, even those presumed to be spread predominantly by the droplet mode. The targeted use of a respirator was also effective, whereas no efficacy was demonstrated for medical masks alone. However, the trends suggest some degree of protection from medical masks, and larger studies are required to measure the efficacy of these devices. The superiority of respirators should be reflected in infection control guidelines to ensure the occupational health and safety of HCWs. A growing body of clinical efficacy evidence, including this study, challenges long-held paradigms about the transmission of infection.

6 | SUMMARY OF KEY POINTS

1. The data collected during two similar clinical trials conducted in Beijing, China, which examined the same infection outcomes, were pooled
2. We showed that respirators provide superior protection against droplet-transmitted infections, for which most guidelines recommend masks. These findings challenge the paradigm of infection transmission being simplified to droplet, airborne or contact.
3. For many infections, more than one mode of transmission is possible, and our data suggest that transmission of infections is more complex than suggested by these paradigms.
4. Clinical efficacy data are a higher level of evidence than theoretical paradigms of transmission, and show better protection afforded by respirators.

ACKNOWLEDGEMENTS

Thanks to the staff at the Beijing Centre for Disease Control and hospitals staff in both trials. Thanks to Australian National Health & Medical Research Council of Australia (Grant # 630787). We acknowledge the guidance and support of 3M China for fit testing during China 1 trial and supplies of masks/respirators during China Trial 2 due to shortage during the 2009 influenza pandemic.

COMPETING INTERESTS

All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare that; (i) Professor C. Raina MacIntyre: Raina MacIntyre has held an Australian Research Council Linkage Grant with 3M as the industry partner, for investigator-driven research. 3M have also contributed supplies of masks and respirators for investigator-driven clinical trials. She has received research grants and laboratory testing as in-kind support from Pfizer, GSK and Bio-CSL for investigator-driven research; (ii) Dr Holly Seale had a NHMRC Australian based Public Health Training Fellowship at the time of the study (1012631). She has also received funding from vaccine manufacturers GSK, bio-CSL and Saniofi Pasteur for investigator-driven research and presentations, and (iii) Dr. Abrar Chughtai had testing of filtration of masks by 3M for PhD. The remaining authors declare that they have no competing interests and have no non-financial interests that may be relevant to the submitted work.

AUTHORS' CONTRIBUTIONS

CRM performed the lead investigator, conception and design of the study, analysing the data and manuscript writing; AAC involved in the statistical analysis and drafting of manuscript; BR performed the contribution to the statistical analysis and revision of manuscript; YP, YZ, HS and XW performed the data management and revision of manuscript; QW performed the contribution to design, analysis and revision of study. All authors have read and approved the final version of the manuscript and ensure that this is the case.

TRIAL REGISTRATION

This is pooled analysis of two clinical trials.

1. First clinical trial—Clinical trial registered with Australian New Zealand Clinical Trials Registry (ANZCTR) (http://www.anzctr.org.au), registration number ACTRN 12609000257268, registered on 13/05/2009.
2. Second clinical trial—Clinical trial registered with Australian New Zealand Clinical Trials Registry (ANZCTR) (http://www.anzctr.org.au), registration number ACTRN 12609000778280, registered on 8/09/2009.

ORCID

Abrar Ahmad Chughtai http://orcid.org/0000-0003-4203-7891

REFERENCES

1. Jacobs JL, Ohde S, Takahashi O, et al. Use of surgical face masks to reduce the incidence of the common cold among health care workers in Japan: a randomized controlled trial. *Am J Infect Control*. 2009;37:417-419.
2. Loeb M, Dafoe N, Mahony J, et al. Surgical mask vs N95 respirator for preventing influenza among health care workers: a randomized trial. *JAMA*. 2009;302:1865-1871.
3. MacIntyre CR, Wang Q, Cauchemez S, et al. A cluster randomized clinical trial comparing fit-tested and non-fit-tested N95 respirators to medical masks to prevent respiratory virus infection in health care workers. *Influenza Other Respir Viruses*. 2011;5:170-179.
4. MacIntyre CR, Wang Q, Seale H, et al. A randomized clinical trial of three options for N95 respirators and medical masks in health workers. *Am J Respir Crit Care Med*. 2013;187:960-966.
5. MacIntyre CR, Seale H, Dung TC, et al. A cluster randomised trial of cloth masks compared with medical masks in healthcare workers. *BMJ Open*. 2015;22:1-10.
6. MacIntyre CR, Chughtai AA, Seale H, et al. Respiratory protection for healthcare workers treating Ebola virus disease (EVD): are facemasks sufficient to meet occupational health and safety obligations? *Int J Nurs Stud*. 2014;51:1421-1426.
7. Wells WF, Winslow CE, Robertson EC. Bacteriologic procedures in the evaluation of methods for control of air-borne infection. *Am J Public Health Nation's Health*. 1946;36:324-331.
8. Osterholm MT, Moore KA, Kelley NS, et al. Transmission of ebola viruses: what we know and what we do not know. *MBio*. 2015;6:e00137-15.
9. Brosseau LM, Jones R. Health workers need optimal respiratory protection for Ebola. The Center for Infectious Disease Research and Policy (CIDRAP) 2014. http://www.cidrap.umn.edu/news-perspective/2014/09/commentary-health-workers-need-optimal-respiratory-protection-ebola. Accessed September 29, 2014.
10. Public Health Agency of Canada. Pathogen Safety Data Sheets and Risk Assessment. http://www.phac-aspc.gc.ca/lab-bio/res/psds-ftss/index-eng.php. Accessed November 10, 2014.
11. Center for Disease Control and Prevention (CDC). Respiratory Syncytial Virus Infection (RSV): Transmission and Prevention. http://www.cdc.gov/rsv/about/transmission.html. Accessed November 10, 2014.
12. Public Health Agency of Canada. Respiratory syncytial virus. http://www.phac-aspc.gc.ca/lab-bio/res/psds-ftss/pneumovirus-eng.php. Accessed November 4, 2014.
13. Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synth Methods*. 2012;3:80-97.
14. Center for Disease Control and Prevention (CDC). Prevention Strategies for Seasonal Influenza in Healthcare Settings. 2014. http://www.cdc.gov/flu/professionals/infectioncontrol/healthcaresettings.htm. Accessed August 19, 2014.
15. World Health Organization (WHO). Infection prevention and control of epidemic- and pandemic-prone acute respiratory infections in health care. 2014.
16. Bischoff WE, Swett K, Leng I, et al. Exposure to influenza virus aerosols during routine patient care. *J Infect Dis*. 2013;207:1037-1046.
17. Lindsley WG, Blachere FM, Thewlis RE, et al. Measurements of airborne influenza virus in aerosol particles from human coughs. *PLoS ONE*. 2010;5:e15100.
18. Siegel JD, Rhinehart E, Jackson M, et al. 2007 Guideline for Isolation Precautions: preventing Transmission of Infectious Agents in Health Care Settings. *Am J Infect Control*. 2007;35(10 Suppl 2):S65-S164.
19. Center for Disease Control and Prevention (CDC). Guidance on personal protective equipment to be used by healthcare workers during management of patients with Ebola Virus Disease in U.S. Hospitals, including procedures for putting on (Donning) and removing (Doffing) 2014. http://www.cdc.gov/vhf/ebola/hcp/procedures-for-ppe.html. Accessed October 23, 2014.
20. Center for Disease Control and Prevention (CDC). Interim infection prevention and control recommendations for hospitalized patients with Middle East respiratory syndrome CoronaVirus (MERS-CoV) 2014. http://www.cdc.gov/coronavirus/mers/infection-prevention-control.html. Accessed September 27, 2014.
21. MacIntyre CR, Chughtai AA, Seale H, et al. Uncertainty, risk analysis and change for Ebola personal protective equipment guidelines. *Int J Nurs Stud*. 2014;52:899-903.
22. World Health Organization (WHO). Ebola and Marburg virus disease epidemics: preparedness, alert, control, and evaluation. 2014.
23. World Health Organisation (WHO). Infection prevention and control during health care for probable or confirmed cases of novel coronavirus (nCoV) infection.
24. Centers for Disease Control and Prevention and World Health Organization. Infection control for viral haemorrhagic fevers in the African health care setting. Atlanta, Centers for Disease Control and Prevention. 1998; 1-198.
25. MacIntyre CR, Chughtai AA. Facemasks for the prevention of infection in healthcare and community settings. *BMJ*. 2015;350:h694.
26. Chughtai AA, Seale H, MacIntyre CR. Availability, consistency and evidence-base of policies and guidelines on the use of mask and respirator to protect hospital health care workers: a global analysis. *BMC Res Notes*. 2013;6:216.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: MacIntyre CR, Chughtai AA, Rahman B, et al. The efficacy of medical masks and respirators against respiratory infection in healthcare workers. *Influenza Other Respi Viruses*. 2017;11:511-517. https://doi.org/10.1111/irv.12474

ACP Journals

Letters | 7 July 2020

Annals of Internal Medicine

Effectiveness of Surgical and Cotton Masks in Blocking SARS–CoV-2: A Controlled Comparison in 4 Patients

FREE

Seongman Bae, MD, Min-Chul Kim, MD, Ji Yeun Kim, PhD, Hye-Hee Cha, BS, Joon Seo Lim, PhD, ... View all authors +

Author, Article and Disclosure Information

https://doi.org/10.7326/M20-1342

This article has been retracted. See Notice of Retraction.

Background: During respiratory viral infection, face masks are thought to prevent transmission (1). Whether face masks worn by patients with coronavirus disease 2019 (COVID-19) prevent contamination of the environment is uncertain (2, 3). A previous study reported that surgical masks and N95 masks were equally effective in preventing the dissemination of influenza virus (4), so surgical masks might help prevent transmission of severe acute respiratory syndrome–coronavirus 2 (SARS–CoV-2). However, the SARS–CoV-2 pandemic has contributed to shortages of both N95 and surgical masks, and cotton masks have gained interest as a substitute.

Objective: To evaluate the effectiveness of surgical and cotton masks in filtering SARS–CoV-2.

Methods and Findings: The institutional review boards of 2 hospitals in Seoul, South Korea, approved the protocol, and we invited patients with COVID-19 to participate. After providing informed consent, patients were admitted to negative pressure isolation rooms. We compared disposable surgical masks (180 mm × 90 mm, 3 layers [inner surface mixed with polypropylene and polyethylene, polypropylene filter, and polypropylene outer surface], pleated, bulk packaged in cardboard; KM Dental Mask, KM Healthcare Corp) with reusable 100% cotton masks (160 mm × 135 mm, 2 layers, individually packaged in plastic; Seoulsa).

A petri dish (90 mm × 15 mm) containing 1 mL of viral transport media (sterile phosphate-buffered saline with bovine serum albumin, 0.1%; penicillin, 10 000 U/mL; streptomycin, 10 mg; and amphotericin B, 25 μg) was placed approximately 20 cm from the patients' mouths. Patients were instructed to cough 5 times each onto a petri dish while wearing the following sequence of masks: no mask, surgical mask, cotton mask, and again with no mask. A separate petri dish was used for each of the 5 coughing episodes. Mask surfaces were swabbed with aseptic Dacron swabs in the following sequence: outer surface of surgical mask, inner surface of surgical mask, outer surface of cotton mask, and inner surface of cotton mask.

The median viral loads of nasopharyngeal and saliva samples from the 4 participants were 5.66 log copies/mL and 4.00 log copies/mL, respectively. The median viral loads after coughs without a mask, with a surgical mask, and with a cotton mask were 2.56 log copies/mL, 2.42 log copies/mL, and 1.85

log copies/mL, respectively. All swabs from the outer mask surfaces of the masks were positive for SARS–CoV-2, whereas most swabs from the inner mask surfaces were negative (Table).

Table. SARS–CoV-2 Viral Load in Patient Samples, Petri Dishes, and Mask Surfaces

Table. SARS-CoV-2 Viral Load in Patient Samples, Petri Dishes, and Mask Surfaces

Characteristic	Patient 1 (Hospital A)	Patient 2 (Hospital A)	Patient 3 (Hospital B)	Patient 4 (Hospital B)
Age, *y*	61	62	35	82
Sex	Male	Female	Male	Female
Clinical diagnosis	Pneumonia	Upper respiratory infection	Upper respiratory infection	Pneumonia with ARDS
Symptom onset before admission, *d*	24*	4	5	10
Timing of the mask test, *hospital days*	8	4	2	14
Viral load, *log copies/mL*				
Nasopharyngeal swab	7.68	5.42	5.98	3.57
Saliva	4.29	2.59	5.91	3.51
Petri dish				
Coughing without a mask (before control)	3.53	2.14	2.52	ND
Coughing with a surgical mask	3.26	1.80	2.21	ND
Coughing with a cotton mask	2.27	ND	1.42	ND
Coughing without a mask (after control)	3.23	2.06	2.64	2.44
Mask surface				
Outer surface of surgical mask	2.21	2.11	2.63	2.59
Inner surface of surgical mask	ND	ND	2.00	ND
Outer surface of cotton mask	2.76	2.66	3.61	2.58
Inner surface of cotton mask	ND	ND	3.70	ND

ARDS = acute respiratory distress syndrome; ND = not detected; SARS-CoV-2 = severe acute respiratory syndrome–coronavirus 2.
* Transferred from the other hospital.

Discussion: Neither surgical nor cotton masks effectively filtered SARS–CoV-2 during coughs by infected patients. Prior evidence that surgical masks effectively filtered influenza virus (1) informed recommendations that patients with confirmed or suspected COVID-19 should wear face masks to prevent transmission (2). However, the size and concentrations of SARS–CoV-2 in aerosols generated during coughing are unknown. Oberg and Brousseau (3) demonstrated that surgical masks did not exhibit adequate filter performance against aerosols measuring 0.9, 2.0, and 3.1 μm in diameter. Lee and colleagues (4) showed that particles 0.04 to 0.2 μm can penetrate surgical masks. The size of the SARS–CoV particle from the 2002–2004 outbreak was estimated as 0.08 to 0.14 μm (5); assuming that SARS-CoV-2 has a similar size, surgical masks are unlikely to effectively filter this virus.

Of note, we found greater contamination on the outer than the inner mask surfaces. Although it is possible that virus particles may cross from the inner to the outer surface because of the physical pressure of swabbing, we swabbed the outer surface before the inner surface. The consistent finding of virus on the outer mask surface is unlikely to have been caused by experimental error or artifact. The mask's aerodynamic features may explain this finding. A turbulent jet due to air leakage around the mask edge could contaminate the outer surface. Alternatively, the small aerosols of SARS–CoV-2 generated during a high-velocity cough might penetrate the masks. However, this hypothesis may only be valid if the coughing patients did not exhale any large-sized particles, which would be expected to be deposited on the inner surface despite high velocity. These observations support the importance of hand hygiene after touching the outer surface of masks.

This experiment did not include N95 masks and does not reflect the actual transmission of infection from patients with COVID-19 wearing different types of masks. We do not know whether masks shorten the travel distance of droplets during coughing. Further study is needed to recommend whether face masks decrease transmission of virus from asymptomatic individuals or those with suspected COVID-19 who are not coughing.

In conclusion, both surgical and cotton masks seem to be ineffective in preventing the dissemination of SARS–CoV-2 from the coughs of patients with COVID-19 to the environment and external mask surface.

This article was published at Annals.org on 6 April 2020.

Comments

20 Comments

SIGN IN TO SUBMIT A COMMENT

ken • palmar • 21 May 2020

inhalation of virus

I'm interested in if masks can prevent inhalation of viruses - not so much if one can sneeze or cough thru a mask.

JSmith • None • 19 May 2020

Methodology clarification

Within the portion of the study under methodology it states, “Patients were instructed to cough 5 times each onto a petri dish while wearing the following sequence of masks…”How were the masks brought into the room? Where were they stored during testing prior to each mask being used? How was each mask placed on each patient? In other words, what precautions were taken to ensure the masks were not contaminated prior to placement on the subjects, during placement on the subjects, or during use? Similarly, how did the experiment ensure no other possible sources for transferred viral load? Were the hands of those who placed the masks on the subjects cleaned and then tested for viral load to ensure no cross contamination? Were control masks included that accompanied the test masks and also tested for viral loads with the same procedures?

Paul W Leu • University of Pittsburgh • 19 May 2020

Irrelevent to Efficacy of Masks

The conclusions of this study by Bae et. al are not only erroneous but misleading. 1. The main result of this study is that higher concentrations of SARS-CoV-2 were found on the outside of masks that were coughed into as opposed to the inside. The fact that the virus was determined to be present on the outside of the mask is unsurprising. Surgical and cotton masks are fabrics which will simply absorb any droplets they come into contact with. The higher concentrations found on the outside of the masks may be due to their swabbing the outside of the masks first (which may remove some of the virus) as opposed to the inside. Results should be compared with swabbing the inside first and then the outside. 2. The presence of SARS-CoV-2 on the outside of masks of infected people is of very limited concern for transmission. Most people put on and remove their own masks and do not touch each other’s masks. 3. The results of this study do NOT show that masks are "ineffective in preventing the dissemination of SARS–CoV-2 from the coughs of patients with COVID-19 to the environment.” As the authors acknowledge, their study does NOT evaluate the ability of the masks to shorten the trajectory of droplets emitted during coughing. The function of the mask is to reduce how far aerosol droplets travel during

breathing, speaking, singing, sneezing, or coughing. This is the same reason one should cover one's mouth or nose with your forearm, inside of your elbow, or tissue when sneezing. CDC guidelines advise the wearing of face coverings to "slow the spread of the virus and help people who may have the virus and do not know it from transmitting it to others."

Joe Breuer • none • 19 May 2020

Theory on negative result on inside of masks

This is a layman's idea for a possible explanation of the counterintuitive result that for the most part the outside of masks tested positive and the insides negative.I cannot speak on its validity and just wish to posit it for discussion by experts. If it possibly leads to valuable insights, great; if it's off base, I hope I did not waste anyone's time.How about the patients expel, along with the virus, other material/cells - related to their immune system or not - that inactivates the virus? And this material or cells *cannot* pass through the masks, so on the insides the inactivation continues / takes place, whereas towards the outside only the infectious material is transported.

Ki Ho Hong1, So Yeon Kim2, Jaehyeon Lee3 • Department of Laboratory Medicine, Seoul Medical Center, Seoul1; National Medical Center, Seoul2, Jeonbuk National University Medical School and Hospital, Jeonju3, South Korea • 17 May 2020

More detailed information about the experiment is needed.

We read with interest Bae and colleagues' study of the effectiveness of surgical and cotton masks against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (1). Although their study is important, some points should be clarified prior to drawing conclusions.The first issue is contamination of aerosolized or small particles. Coughing can cause environmental contamination. Nevertheless, all experiments appeared to have been performed within the same room. Coughing into Petri dishes without a mask could have produced airborne virus-containing droplets that contaminated the next steps of experiment. The sequences would be better following; surgical mask, cotton mask, and no mask. Additionally, fitted mask are critical for preventing room contamination, but mask fit was not discussed (2).

Second, swabbing the inner surfaces of masks may not have been sufficient for the "not detected" results in their Table (1). Removing mask layer and subsequent particle elution in media for nucleic acid amplification might have been a better alternative (3), especially for the evaluation of inner surfaces. Third, the results and conclusions appeared to differ. Cotton masks reduced virus titers by 1-2 log10 copies/mL in Patients 1, 2, and 3. For Patient 4, the virus was detected only in the coughing-without-a-mask Petri dish and only on the outer surfaces of her surgical and cotton masks.

The conclusion that cotton masks do not effectively filter SARS-CoV-2 does not correspond to these findings. In addition, the number of experiments was too small for the conclusion.Finally, a few PCR results in the report (1) were under the measurable level by most PCR protocols widely used. The authors did not describe the PCR protocol adopted or the analytical performance of it with the limit of detection (LoD). Most sensitive LoD theoretically possible is 3 copies/reaction. Assuming that the total reaction volume and RNA volume for the PCR reaction were 5 and 25 μL, respectively, and that the widely used QIAamp Viral RNA MINI kit (Qiagen,

Hilden, Germany) was utilized for RNA extraction per the manufacturer's protocol, the LoD should be approximately 2.41 log copies/mL. Indeed, Corman reported 2.31 log copies/mL based on a 25 μL reaction volume (4), and Pfefferle reported 2.83 log copies/mL (5). Additionally, the limit of quantitation is usually higher than the LoD. The authors reported 1.42 log10 copies/mL, which appears too low. Results below LoD should be reported as "less than LoD." Before resolving these issues, the conclusion should be interpreted with caution.

References

1.Bae S, Kim M-C, Kim JY, et al. Effectiveness of surgical and cotton masks in blocking SARS-CoV-2: a controlled comparison in 4 Patients. Ann Int Med. 2020. [Epub ahead of print]. PubMed PMID: 32251511. doi:10.7326/M20-1342

2.Patel RB, Skaria SD, Mansour MM, et al. Respiratory source control using a surgical mask: an in vitro study. J Occup Environ Hyg. 2016;13:569–576. PubMed PMID: 26225807.

3.Chughtai AA, Stelzer-Braid S, Rawlinson W, et al. Contamination by respiratory viruses on outer surface of medical masks used by hospital healthcare workers. BMC Infect Dis. 2019;19(1):491. PubMed PMID: 31159777.

4.Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. 2020;25. PubMed PMID: 31992387.5.Pfefferle S, Reucher S, Nörz D, et al. Evaluation of a quantitative RT-PCR assay for the detection of the emerging coronavirus SARS-CoV-2 using a high throughput system. Euro Surveill. 2020;25.PubMed PMID: 32156329.

AcknowledgmentsAuthor contributionsDrs. Ki Ho Hong and So Yeon Kim contributed equally as co-first authors. Drs. Jaehyeon Lee and Ki Ho Hong drafted the manuscript. All authors participated in the concept development and critical revision of the manuscript for important intellectual content.

Conflicts of interest All authors declare no conflict of interest.

Angela C • None • 16 May 2020

Need a larger sample size and more control tests

Hello,I am very interested in this study, as I have my own suspicions about the effectiveness of wearing face masks. I would like to see a larger, more random study conducted. I don't feel a study with 4 participants can give you reliable data. Also, I would like to see the inside and outside of these masks tested prior to the cough tests being completed. And finally, I didn't see the results of the viral load on each of the Petri dishes that the participants coughed onto. That would be interesting to have them cough into Petri dishes at various distances wearing different face masks, including the N95. Thank you for starting this study. I think it needs more work though. I have been a paid RN, BSN for the past 20 years, but my opinions do not reflect those of my employers.

Sung-Han Kim, MD. • Asan Medical Center • 28 April 2020

Author's response

We totally agree with Dr. Glele and colleagues' comment on the high variability of coughing intensity within subjects. Furthermore, it is worth to note that one of eight coughing experiments without mask in patients with COVID-19 revealed a negative SARS-CoV-2 PCR result (Table 1). The heterogeneity of transmission of coronavirus including SARS-CoV, MERS-CoV, and SARS-CoV-2 may explain this observation. The recent study reported that none 41 healthcare workers with most surgical masks and minor N95 masks who were exposed to the aerosol-generating procedures in eventually diagnosed COVID-19 patients developed symptoms, and all PCR tests for SARS-CoV-2 were negative (COVID-19 and the risk to health care workers: a case report. Ann Intern Med 2020 March 16). Given that viral expectoration from coughing COVID-19 patients was not uniform based on our experiment, cautious interpretation for unusual transmission events is always needed. Dr. Glele and colleagues also commented that no detection of SARS-CoV-2 RNA from inner surface except one patient precludes any reliable conclusions. We assume that multiple factors may affect swab sampling from the outer and inner surfaces of the masks. Although environmental sampling from hard surfaces such as plastic or metal has been widely studied, there are limited studies on sampling from fabric materials. Elution of punched layers of face masks may provide more valuable information about the surface contamination of the masks. Further studies are needed on the viral contamination of mask surfaces. In this context, this variability of viral shedding from coughing within the subject and the nature of fabric swab sampling should be bear in mind for the interpretation of our small experimental data.

As Dr. Glele and colleagues' comment, Leung et al. reported the efficacy of surgical masks in reducing coronavirus detection and viral load from 17 patients (Nat Med 2020 Apr 3). The big difference between Leung's study and ours is the method of collecting human coronavirus particles from the patients. Leung's study collected virus particles by a closed system such as G-II bioaerosol collecting device which consists of a large cone connected with a closed duct. In contrast, we collected virus particle of SARS-CoV-2 directly from coughing COVID-19 patients with an open air system in a negative pressure room. Furthermore, the results of the efficacy of surgical masks on influenza virus from Leung's study (Nat Med 2020 Apr 3) are different from those by the previous study (Clin Infect Dis 2009; 49:275-7). The different methodology of sample collection may explain this discrepancy.

Dr. Purens and colleagues pointed the statistical issue. Our complete case analysis (CCA) may overestimate the true value. In contrast, if we included "not detectable" as "zero", the calculation may underestimate the true value. So, an alternative calculation such as single imputation or Dr. Purens' calculation may result in the value between these two. Thank you for suggesting one of good sensitivity analysis.

We appreciated Dr. Yeung's good balanced view of our study results. We agree with Dr. Yeung's opinion on that our small study (n=4) is a pilot study. We have recently completed additional mask tests in 7 COVID-19 patients to compare the use of surgical masks to the use of N95-equivalent respirators. We believe that these data will provide more information on this issue. Furthermore, other independent groups should evaluate the outward and inward protective effectiveness of various masks against SARS-CoV-2 with more well-designed protocols in which the issues raised in this pilot study by many experts

can be settled. Therefore, we totally agree with Dr. Yeung's view on this pilot study like the glass half full or empty.

Christopher T. Leffler, MD, MPH.1 Edsel Ing MD, MPH, CPH, MIAD.2 Joseph D. Lykins V, MD.1 Craig A. McKeown, MD3. Andrzej Grzybowski, MD.4 • 1. Virginia Commonwealth University 2. University of Toronto 3. University of Miami 4. University of Warmia and Mazury • 30 April 2020

Prevention of the spread of coronavirus using masks.

We read the work which concluded "both surgical and cotton masks seem to be ineffective in preventing the dissemination of SARS–CoV-2…"1 In fact, compared with the control condition, the petri dish viral load was less with a cloth mask for all patients, and in half, was not detectable.1

Such reductions do help at the population level.2,3 We retrieved mortality and testing data for 169 countries from a publicly available source on April 22, 2020.4 On average, the time from infection to symptoms is 5.1 days, and that from infection to death is 23 days.2 Therefore, the date of each country's initial infection was estimated as the earlier of: 5 days before the first reported infection, or 23 days before the first death.4,5 As deaths by April 22, 2020 would typically reflect infections beginning 23 days previously (by March 30), both the time from the first infection, and from the time the public began wearing masks, until March 30 were determined. Countries in which mask usage has been widespread include Hong Kong, South Korea, Malaysia, Taiwan, Japan, and Mongolia.2 Mandates for wearing of masks in public had been issued by March 30 in Thailand (March 12), Vietnam (March 16), Czechia (March 19), and Slovakia (March 25).2 The exponential growth associated with the spread of an epidemic appears linear on a logarithmic scale.2 By multivariable linear regression, significant predictors of the logarithm of each country's per-capita coronavirus mortality included: duration of infection in the country, duration of wearing masks, population size, and per-capita testing (all $p<0.001$, Table 1). In a population not wearing masks, the per-capita mortality tended to increase each week by a factor of $10^{0.156} = 1.43$, or 43%. On the other hand, in a population wearing masks, the per-capita mortality tended to increase by a factor of $10^{(0.156-0.144)} = 1.028$, or just 2.8%. The positive association with testing probably reflects the greater recognition of coronavirus-related mortality with more testing, as well as the increased incentive countries have to test when they suffer a more intense outbreak. These results support the universal wearing of masks by the public to suppress the spread of the coronavirus. Mask-wearing should be adopted immediately, based on the precautionary principle.2,3

References.

1. Bae S, Kim MC, Kim JY, Cha HH, Lim JS, Jung J, Kim MJ, Oh DK, Lee MK, Choi SH, Sung M. Effectiveness of surgical and cotton masks in blocking SARS–CoV-2: a controlled comparison in 4 patients. Annals of Internal Medicine. 2020 Apr 6.

2. Leffler CT, Ing E, McKeown CA, Pratt D, Grzybowski A. Final Country-wide Mortality from the Novel Coronavirus (COVID-19) Pandemic and Notes Regarding Mask Usage by the Public. April 4, 2020. Available from: https://www.researchgate.net/publication/340438732_Country-wide_Mortality_from_the_Novel_Coronavirus_COVID-

19_Pandemic_and_Notes_Regarding_Mask_Usage_by_the_Public DOI: 10.13140/RG.2.2.36006.27200
3. Howard J, Huang A, Li Z, Tufekci Z, et al. Face masks against COVID-19: an evidence review. Preprints 2020; published online April 12. DOI:10.20944/preprints202004.0203.v1 (preprint).
4. Worldometers. COVID-19 Coronavirus Pandemic. Available from: https://www.worldometers.info/coronavirus/?utm_campaign=homeAdUOA?Si Accessed April 22, 2020.
5. European Centre for Disease Prevention and Control. COVID-19 Coronavirus data. Available from: https://data.europa.eu/euodp/en/data/dataset/covid-19-coronavirus-data
Accessed April 16, 2020.

None of the authors has any conflicts of interest to disclose.

Table 1. Predictors of (log) Country-wide Per-capita Coronavirus Mortality by Multivariable Linear Regression in 169 Countries.

Coefficient (SE) 95% CI P value.

Duration in country (weeks) 0.156 (SE 0.034) (95% CI 0.089 to 0.223) p<0.001.

Time wearing masks (weeks) -0.144 (SE 0.033) (95% CI -0.209 to -0.079) p<0.001.

Population (log) -0.297 (SE 0.079) (95% CI -0.453 to -0.141) p<0.001.

Tests per capita (log) 0.612 (SE 0.085) (95% CI 0.445 to 0.779) p <0.001.

Constant -2.571 (SE 0.368) (95% CI -3.299 to -1.844) p<0.001.

Eugene Y.H. Yeung • Faculty of Medicine, University of Ottawa; Eastern Ontario Regional Laboratory Association (EORLA) • 27 April 2020

Effectiveness of Masks in Blocking SARS-CoV-2: Depends on Whether You See the Glass Half Full or Empty

It is difficult to draw a solid conclusion from a study of 4 participants, which clearly lacked statistical power to detect difference between control and intervention groups. This is a pilot study at best, but our interpretation depends on whether we see the glass half full or empty. Optimistic researchers would notice a trend of decrease in SARS-CoV-2 viral load when each participant had face mask on. Although the study found contamination on the outer surface of face masks, there was no evidence that the viral particles bypassed the mask and entered the wearers' mucosa. Three of the 4 participants had undetectable viral load in inner surface of masks. These findings suggested potential role of masks as barriers against entrance of viral particles. Optimistic researchers would be satisfied with these preliminary findings, and thereby conduct a larger study with sufficient statistical power. On the contrary,

pessimistic researchers would see this study as a failure and conclude masks are ineffective in preventing the dissemination of SARS–CoV-2. As Sir Winston Churchill stated, "A pessimist sees the difficulty in every opportunity; an optimist sees the opportunity in every difficulty."

Disclosures: I have been paid for working in primary and secondary care settings, but not for writing this letter. Opinions expressed are solely my own and do not express the views of my employer.

Ludwig Serge Aho Glele, Sara romano-Bertrand, Jean-Francois Gehanno, Didier Lepelletier • Epidemiology, infection control, evaluation. Dijon, Montpellier, Rouen, Nantes. France and public health • 27 April 2020

General response to Bae et al.

We read with interest the article by Seongman Bae et al. (1) estimating the blocking power of surgical mask and cotton mask against SARS-CoV-2.
Patients with known viral loads had to cough five times in a petri dish following the sequence: no mask, surgical mask, cotton mask then no mask again. Different petri dishes were used for each of the five cough episodes and we assume that each patient coughed 5 times on each petri dish for each step of the sequence, as there were only four steps by sequence.
Authors implicitly consider that the intensity of coughing does not vary between subjects and during the course of the experiment, which is not in line with the high variability within subjects (2).
Outcomes criteria were the contamination of petri dishes, and of external and internal surfaces of masks. No air samples were collected close to patients along with the experiment but it would be informative on SARS-CoV-2 shedding through ineffective masks.
Outer surfaces of masks were more contaminated than inner surfaces, but this was in fact assessed only for one patient (patient 3), since inner surface contamination was not detected for the three other patients. This precludes any statistical test and therefore any reliable conclusion.
Authors based the statement that neither surgical nor cotton masks effectively filtered SARS–CoV-2 during coughs on only two patients (1 and 3) without any statistical test. The median viral loads (log copies/mL) in nasopharynx and saliva from the four participants were respectively of 5.66 and 4.00, but varied from 3.51 to 7.68. Furthermore viral loads, when detected, were often very close to the RT-PCR detection limit. This can induce bias but is not discussed by authors. We therefore consider that their statement cannot be considered reliable. A study on 17 patients demonstrated the efficacy of surgical masks in reducing coronavirus detection and viral loads in both large respiratory droplets and aerosols (3). Non-parametric tests can be performed even with very small samples (4). Potential confounding factors, particularly viral loads, were collected but were not statistically analysed. Larger sample size would have allowed the development of an experimental design that could consider: initial viral load level and correlation of the data (difference in viral load between outer and inner surfaces, initial level in the oropharynx and mask contamination, contamination of petri dishes and surfaces...). Such a more complex experimental design (5) would allow more reliable conclusions.

References

1. Bae S, Kim M-C, Kim JY, Cha H-H, Lim JS, Jung J, et al. Effectiveness of Surgical and Cotton Masks in

Blocking SARS-CoV-2: A Controlled Comparison in 4 Patients. Ann Intern Med. 2020 Apr 6;
2. Duguid JP. The size and the duration of air-carriage of respiratory droplets and droplet-nuclei. J Hyg (Lond). 1946 Sep;44(6):471–9.
3. Leung NHL, Chu DKW, Shiu EYC, Chan K-H, McDevitt JJ, Hau BJP, et al. Respiratory virus shedding in exhaled breath and efficacy of face masks. Nat Med [Internet]. 2020 Apr 3 [cited 2020 Apr 6]; Available from: http://www.nature.com/articles/s41591-020-0843-2
4. Dwivedi AK, Mallawaarachchi I, Alvarado LA. Analysis of small sample size studies using nonparametric bootstrap test with pooled resampling method. Statistics in Medicine. 2017;36(14):2187–205.
5. Baker TB, Smith SS, Bolt DM, Loh W-Y, Mermelstein R, Fiore MC, et al. Implementing Clinical Research Using Factorial Designs: A Primer. Behav Ther. 2017;48(4):567–80.

Ken Lim • CyberMedia Convergence consulting • 20 April 2020

Principal Investigator

Extremely flawed experiment got published with n=4!!! OFC the virus went thru! the masks aren't waterproof. OFC it went thru, force of cough pushed it thru. The test should be how many viral droplets appeared on another person or surface 2-3m away! The test should be hi-speed video showing # of particles expelled w & w/o mask! Unbelievably poor experiment!

Kristopher Purens, PhD, Abigail Purens, DVM/MPH candidate • Descartes Labs, Inc., University of Minnesota College of Veterinary Medicine • 17 April 2020

Statistical analysis shows decreased airborne SARS-CoV-2 transmission with the use of masks in line with previous studies

To test the efficacy of masks to reduce respiratory transmission of SARS-CoV-2, Bae et al.(1) replicated methods previously published by Johnson et al. (2009)(2) in an important early comparative study. A precautionary approach to new public health threats such as the COVID-19 pandemic is to use the best available models as analogues, make conservative recommendations, and update as new data become available. This necessitates careful null hypothesis selection and an information-gained approach to new data and ongoing analysis. A precautionary null hypothesis to COVID-19 is to test whether new evidence is strong enough to reject prior recommendations, such as widespread mask use. Johnson found that masks reduced respiratory transmission of influenza virus, a disease commonly used as a model for SARS-CoV-2.(3) In this context, Bae's null hypothesis that masks do not reduce viral load transmission was inappropriate. Combined with Bae's small sample size, this led to reporting of mask wearing causing no significant reduction in SARS-CoV-2 viral load transmission, in contrast to Johnson's findings for influenza.

Additionally, statistical analyses for non-normally distributed data and small sample size are appropriate in this context, to prevent being misled by violating the assumptions of common statistical methods. Two such appropriate analyses are probability based methods, and permutation tests. Analytical power can be increased by treating each pair of masked/non-masked attempts as a trial, and correcting for differences in base viral load for each individual.(4) We assumed no detection (ND) just below the lowest detected

threshold reported, with differences calculated from that highest-reasonable viral load that would result in ND.

To this end we performed two tests: 1) non-parametric probabilistic approach testing whether Bae's results indicate masks caused no reduction in respiratory SARS-CoV-2 transmission and 2) permutation resampling testing of whether Bae's results were significantly different than Johnson's influenza virus transmission results.(4) Our analysis found that masks provide >0 reduction in viral load transmission ($p=0.0078$) and that Bae's results for SARS-CoV-2 were not significantly different from Johnson's results for influenza in reducing respiratory viral load transmission ($p = 0.158$). Our results support the continued use of influenza as a model for public health decisions regarding SARS-CoV-2. Importantly for public health, our analysis supports current recommendations for widespread mask wearing during the COVID-19 pandemic.(5)

The combined data set assembled, Bae et al. and Johnson et al., and analysis is available at https://github.com/purens/sars_cov2_masks to allow further study.

References

1. Bae S, Kim M-C, Kim JY, Cha H-H, Lim JS, Jung J, et al. Effectiveness of Surgical and Cotton Masks in Blocking SARS–CoV-2: A Controlled Comparison in 4 Patients. Annals of Internal Medicine [Internet]. 2020 Apr 6 [cited 2020 Apr 7]; Available from: https://doi.org/10.7326/M20-1342

2. Johnson DF, Druce JD, Birch C, Grayson ML. A Quantitative Assessment of the Efficacy of Surgical and N95 Masks to Filter Influenza Virus in Patients with Acute Influenza Infection. Clin Infect Dis. 2009 Jul 15;49(2):275–7.

3. Cowling BJ, Ali ST, Ng TWY, Tsang TK, Li JCM, Fong MW, et al. Impact assessment of non-pharmaceutical interventions against COVID-19 and influenza in Hong Kong: an observational study. medRxiv. 2020 Mar 16;2020.03.12.20034660.

4. Sokal RR, Rohlf FJ. Biometry : the principles and practice of statistics in biological research. 4th ed. New York: WH Freeman and Company; 2013. 937 p.

5. CDC. Recommendation Regarding the Use of Cloth Face Coverings, Especially in Areas of Significant Community-Based Transmission [Internet]. Centers for Disease Control and Prevention. 2020 [cited 2020 Apr 5]. Available from: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/cloth-face-cover.html

Sung-Han Kim, MD. • Asan Medical Center • 15 April 2020

Author's response to the comments

Thank you for your thoughtful comments on the concept of the strong ability of airborne transmission of SARS-CoV-2 (Dr. Shu Yuan). So, they commented the possibility of the environmental air contamination before the patients wore the masks. I also agree with that environmental air contamination may result in outer surface contamination of masks and petri dish in front of the patients. Initially, we did not concern about air contamination by coughing without mask, like the previous study (Clin Infect Dis 2009; 49:275-7). So, we performed testing sequence as coughing without mask first. But, it is possible that initial coughing without masks might contaminate the surrounding air, given that NEJM paper demonstrated air stability of SARS-CoV-2. However, negative pressure room where the patients stayed had more than 12 air change per hour, so theoretically 99% of particles is cleared within 23 min. In addition, we used small petri dish, so it is unlikely that aerosol landed on this area of small petri dish during the testing with subsequent mask changes. Actually, we performed air sampling before this experiment to investigate the aerosol transmission in the patients' room. We had collected about 1,000 L air for 20 min by air sampler (Sartorius) like our previous study in MERS infected patients' room (Clin Infect Dis 2016; 63:363-9). We can found a few positive PCR results from air sampling, although we collected air sampling without active coughing (unpublished data). Instead, we assume that fine aerosols leaked from the masks may contaminated the outer surface of the masks. In addition, we hypothesized the spit without virus particle might be deposited in the inner surface of the mask like Dr. Hoehn's comment. However, further well-controlled study with air sampling and more cautious coughing sequence in different rooms may provide us valuable information for these hypotheses.

Dr. Lasica and Dr. Ing commented the statistical points. But, we think that the numerical data presented in this small study do not have any statistical meaning. So, the interpretation based on the median or mean values with the calculation of p value may be not useful. A more adequate powered studies are urgently needed.

Dr. Harada commented that the value for the mask surface is difficult to express at per mL. We used dacron swabs premoistened with viral transport media (3 mL) to swab the outer and inner surfaces of the mask aseptically. So, we expressed the values as per mL.

Dr. Rzymski's comment provide valuable information to us for the designing of further experiments. He suggested that prolonged speaking may be associated with the release of the higher number of droplets than coughing. So, we are now planning to evaluate the efficacies of various types of masks during talking.

We totally agree with Dr. Camioli's comments indicating that there are no evidence about that surgical masks are ineffective for healthcare workers. In addition, we agree with his opinion that masks may reduce the forward momentum of the virus-spit particles. Our small study did not show surgical or cottom masks have no role to spread SARS-CoV-2 to the environment. We assume that surgical mask may be not equivalent to N95-equivalent high efficient masks for outward spreading especially in coughing COVID-19 patients, while we just completed additional experiment using N95-equivalent masks. We did not show that any kind of masks such as cotton or surgical masks have no role to quantitatively reduce the spread of coughing SARS-CoV-2 to the environment. Based on empirical evidence, masks might shorten the distance of aerosol containing virus (Dr. Camioli's comments), redirect the turbulent jets in less harmful directions (outward proection), and reduce the amount of virus particles

from the patients, although the targeted studies using SARS-CoV-2 are lacking. Furthermore, the inhaled air might have different aerodynamics in terms of low velocity particles with adherence of masks to face by depressurizing. So, the ineffectiveness of outward protection of surgical or cotton masks in coughing COVID-19 patients do not mean ineffectiveness inward protection of these masks. As Dr. Camioli's comment and the CDC guidelines, wearing any kind of masks in public settings with hand hygiene is highly recommended.

Cristina Corsini Campioli MD, Stacey Rizza MD, Abinash Virk MD, John C. O'Horo, MD, MPH • Mayo Clinic Rochester, Minnesota • 14 April 2020

Masking in COVID-19: Teach the Controversy

TO THE EDITOR:

The paper by Seongman Bae (1) and colleagues' study regarding the effectiveness of surgical and cotton masks in blocking SARS-CoV-2 presented several unexpected findings. Seongman Bae et al evaluated the amount of virus coughed through a surgical or cotton mask at a distance close to 8 inches in four patients. Virus was recovered at this distance, but more surprisingly, virus was identified on the outer surface of the masks, but not on the inner surface after coughing. The authors conclude that surgical and cotton masks are ineffective at preventing the dissemination of SARS-CoV-2. This is likely to aggravate ongoing controversy regarding personal protective equipment (PPE).

Public health authorities define a significant exposure to SARS-CoV-2 as face-to-face (unmasked) contact within 6 feet with a patient with symptomatic infection. The situation where both a healthcare worker and a patient is masked, as currently recommended by the Centers for Disease Control and Prevention's Universal Masking guideline, was not evaluated in this study. Masks may reduce the forward momentum of the virus-spit particles so that they are not launched as far forward as an unconstrained cough. Testing at a distance of only 8 inches in four patients provides inadequate evidence to stop using these masks for this purpose. The finding of lower viral load on the petri dish compared to the surgical mask goes against the known poor filterability of 2-ply cotton masks. A previous study showed that 2-ply cotton masks are ineffective in preventing respiratory viral infections (RVI) (2), while other studies have demonstrated efficacy of the medical masks in decreasing RVI (3, 4).

This also should not be construed as evidence that surgical masks are ineffective for healthcare workers. Testing how much virus escaped from five coughs is not representative of the effectiveness of these masks at filtering virus during normal respiration. Indeed, a case report in the Annals last month indicated that wearing a surgical mask was adequate PPE for exposure of 41 healthcare workers to a series of aerosol generating procedures in a COVID-19 positive patient (5).

The contribution of this paper is recognizing the significant contamination of the outer surface after coughing. Masking alone without the combination of meticulous hand hygiene, proper doffing and physical distancing, may risk spread of SARS-CoV-2. This article should not be interpreted as advice to the public to forgo masks or evidence against droplet precautions effectiveness for healthcare workers.

1. Bae S, Kim MC, Kim JY, Cha HH, Lim JS, Jung J, et al. Effectiveness of Surgical and Cotton Masks in

Blocking SARS-CoV-2: A Controlled Comparison in 4 Patients. Ann Intern Med. 2020.
2. MacIntyre CR, Seale H, Dung TC, Hien NT, Nga PT, Chughtai AA, et al. A cluster randomised trial of cloth masks compared with medical masks in healthcare workers. BMJ Open. 2015;5(4):e006577.
3. Sung AD, Sung JAM, Thomas S, Hyslop T, Gasparetto C, Long G, et al. Universal Mask Usage for Reduction of Respiratory Viral Infections After Stem Cell Transplant: A Prospective Trial. Clin Infect Dis. 2016;63(8):999-1006.
4. Yeo KT, Yung CF, Chiew LC, Yunus HM, Thoon KC, Gomez M, et al. Universal Mask Policy in the Neonatal Unit to Reduce Respiratory Viral Infections. Clin Infect Dis. 2017;64(6):817.
5. Ng K, Poon BH, Kiat Puar TH, Shan Quah JL, Loh WJ, Wong YJ, et al. COVID-19 and the Risk to Health Care Workers: A Case Report. Ann Intern Med. 2020.

Lukasz Szarpak, Krzysztof J. Filipiak, Milosz Jaguszewski, Jerzy R. Ladny, Jacek Smereka • Lazarski University, Medical University of: Warsaw, Gdansk, Bialystok and Wroclaw • 11 April 2020

Does the use of surgical or cotton masks reduce the risk of SARS-CoV-2 infection?

We have read with great interest the article Bae et al. regarding the effectiveness of the use of surgical and cotton masks in blocking SARS-CoV-2.

This is an important contribution to the discussion on the prevention of SARS-CoV-2 pandemic infections, especially at a time when there is a widespread lack of basic personal protective equipment for medical personnel and other persons exposed to potentially infected or confirmed COVID-19 individuals The rationale for using surgical and cotton masks by potentially healthy persons to reduce transmission of the infection from asymptomatic persons is currently being discussed. Many studies have shown that the effectiveness of medical masks and N95 respirators in reducing the risk of respiratory infections was comparable.

However, in the context of the studies carried out by Bae et al. it should be taken into account that a petri dish containing viral transport media was placed approximately 20 cm from the patients' mouths. Such a short distance was indeed necessary for methodological reasons, however, the results do not indicate the possibility of spreading the aerosol over longer distances and it is still possible that both surgical and cotton masks limit the range of the aerosol with SARS-CoV-2 virus.

The authors in the conclusion stated that surgical and cotton masks seem to be ineffective in preventing the dissemination of SARS-CoV-2 from the coughs of patients with COVID-19 to the environment and external mask surface, but this statement should be complemented by a clear declaration that the samples were taken at a distance of only 20 cm and that these test results do not refer to the possibility of reducing infections.

Lukasz Szarpak, Krzysztof J. Filipiak, Milosz Jaguszewski, Jerzy R. Ladny, Jacek Smereka • Lazarski University, Medical University if Warsaw, Medical University of Gdansk, Medical University of Bialystok, Wroclaw Medical University • 11 April 2020

The use of personal protective equipment in the COVID-19 pandemic era

The current pandemic is reducing medical resources and requires PPE adaptation to the circumstances and to the scale of the threat to medical personnel. One should remember that it is the most important to follow the general recommendations on hand disinfection and the sequence of procedures when putting on and taking off PPE. It is essential to use masks with a filter, but also goggles and visors to protect the face, as well as double or triple gloves (Figure). Sterile surgical gloves are particularly useful as they are longer.

The optimal solution is to fully protect the entire body surface, isolate it from the environment, and breathe in air from a portable source, but this is not necessary in the case of SARS-CoV-2. At present, it is recommended to apply various types of equipment, including, in particular, partial protection of the environment through the use of surgical masks or ordinary face masks by persons with confirmed or potential SARS-CoV-2 infection; this may reduce the risk of infecting people in the environment, including medical personnel.

At present, performing a number of procedures in emergency medicine is associated with additional problems and risks for medical personnel. Emergency physicians, anesthesiologists and intensive care specialists, as well as the relevant scientific societies issue recommendations concerning endotracheal intubation or other procedures dangerous for the medical personnel. It should be remembered that endotracheal intubation by using direct laryngoscopy without adequate protection presents a high risk of SARS-CoV-2 infection. The proposed modifications of endotracheal intubation include special preparation of the equipment and medical personnel, using a special protective box, foils applied to the upper half of the patient's body, and the use of indirect laryngoscopy methods, including video laryngoscopy and rapid sequence intubation. In this context, it should be emphasized that attempts of prehospital endotracheal intubation by inexperienced personnel constitute a challenge, and supraglottic methods should be kept in mind. If intravenous access cannot be established or is technically difficult, it is still possible to establish intraosseous access. Performing several procedures in protective clothing is technically difficult and exhausting, which is especially true for CPR. Certain intra-hospital procedures must be modified, for example, cardiopulmonary resuscitation in a patient with ARDS in a prone position and electrical defibrillation.

The COVID-19 pandemic poses a huge challenge for emergency teams, as well as physicians in emergency departments. The need for additional protection of the patient and medical personnel may result in a significant delay in the arrival of the emergency team, patient transport, and provision of intended medical care. During any pandemic, people still suffer from various diseases and injuries that require treatment. The need to regroup medical forces and resources should not increase morbidity or mortality from diseases other than COVID-19.

Kouji H. Harada, Mariko Harada Sassa • Kyoto University • 11 April 2020

Concerns on the method and data presentation

We express concerns over the values presented in this report. The concerns come from the improper description of the method and findings. Viral loads are described as log copies/mL, but it is not clear to evaluate the results. Particularly, the value for the mask surface is difficult to be expressed at per mL. In addition, detectable level of viral loads in each media is not provided in the report. When comparing

different media, it is inappropriate to simply describe the levels because the amount of the sample and the detection limit are different among media. We are worried about the probable confusion caused by the report.

Disclosures: None.

Piotr Rzymski • Department of Environmental Medicine, Poznan University of Medical Sciences, Poland • 10 April 2020

Effectiveness during speech and normal breathing

It would of high interest and value to conduct a similar study in which the effectiveness of surgical and cotton masks in blocking SARS-CoV-2 is assessed during normal speech. Speaking (as demonstrated by counting to 100) can be associated with the release of the higher number of droplets than a single cough [1, 2], and the rate of emission is related to loudness [3] although the released particles are smaller. Therefore, the force exerted on the mask and associated aerosol penetration should both be lower than in the case of coughing. On the other hand, prolonged speaking in the mask could damp it and eventually lead to the release of droplets containing an infectious agent. Moreover, it would be valuable to investigate whether droplets released during normal breathing by a positive patient can lead to aerosol penetration of a mask and the spread of the virus. Some works have shown that normal breathing, without coughing or sneezing, by influenza-positive patients can lead to the generation of small droplets containing significant number of influenza RNA [4]
Testing the above experimentally would provide some indirect information on whether surgical and cotton masks can be effective in decreased the transmission of the virus before the symptoms are onset. Obviously, it would be best to perform such a study on positive subjects not presenting COVID-19 symptoms although it would be logistically challenging.

[1] Loudon, R. G. & Roberts, R. M. (1968) Droplet expulsion from the respiratory tract. American Review of Respiratory Disease 95, 435–442.
[2] Papineni, R. S. & Rosenthal, F. S. (1997) The size distribution of droplets in the exhaled breath of healthy human subjects. Journal of Aerosol Medicine and Pulmonary Drug Delivery 10, 105–116.
[3] Asadi, S. et al. (2019) Aerosol emission and superemission during human speech increase with voice loudness. Scientific Reports 9, 2348.
[4] Yan, J. et al. (2018) Infectious virus in exhaled breath of symptomatic seasonal influenza cases from a college community. Proceedings of the National Academy of Sciences of the United States of America 115, 1081–1086.

Francisco Bracho • Ventura County Med Ctr, Childrens Hosp Los Angeles Med Group • 9 April 2020

Inside of mask negative?

It looks like a translation error but the inside of the mask could not be negative and the outside positive.

Michal Lasica, PhD • Institute of Mathematics of the Polish Academy of Sciences • 9 April 2020

Apparent serious error in analysis and interpretation of the data

I write as a professional mathematician and a concerned member of the public. Admittedly, I have no professional background in life sciences. However, I see an important flaw in the paper, which seems serious, and may even negate the final conclusion, as it is stated. My concerns were essentially stated by Dr Michael J DeWeert, but I would like to reiterate with more detail.

According to included table, when coughing onto a Petri dish without a barrier, the 4 patients release detectable viral load. When coughing through a cotton mask, in 2 cases the viral load is not detectable (ND), and in the other 2 it is reduced more than 10 times. Yet, according to the average (the authors use the word "median", while they actually compute averages) viral loads presented by the authors as main results, the viral load is reduced only 5 times. This is apparently because in the computations, the averages are taken over whole rows of the table with the ND instances ignored. This is a serious methodological error. If the virus was not detected in 3 patients instead of 2, the average could have been even higher.

As Dr DeWeert stated, this seems to undermine the conclusion that "cotton masks seem to be ineffective in preventing the dissemination of SARS–CoV-2 from the coughs of patients with COVID-19 to the environment". In fact, if a larger-scale study of this kind yielded similar results, this could be a strong argument for the use of cotton masks by general public in advanced stages of the pandemic.
I am particularly concerned that the paper might discourage the use of masks by the public. In fact I learned about the study from an article on a Polish news website, which cited the conclusion of the authors together with the erroneous averages.

PREVIOUS ARTICLE NEXT ARTICLE

BRIEF COMMUNICATION
https://doi.org/10.1038/s41591-020-0843-2

nature medicine

Respiratory virus shedding in exhaled breath and efficacy of face masks

Nancy H. L. Leung[1], Daniel K. W. Chu[1], Eunice Y. C. Shiu[1], Kwok-Hung Chan[2], James J. McDevitt[3], Benien J. P. Hau[1,4], Hui-Ling Yen[1], Yuguo Li[5], Dennis K. M. Ip[1], J. S. Malik Peiris[1], Wing-Hong Seto[1,6], Gabriel M. Leung[1], Donald K. Milton[7,8] and Benjamin J. Cowling[1,8] ✉

We identified seasonal human coronaviruses, influenza viruses and rhinoviruses in exhaled breath and coughs of children and adults with acute respiratory illness. Surgical face masks significantly reduced detection of influenza virus RNA in respiratory droplets and coronavirus RNA in aerosols, with a trend toward reduced detection of coronavirus RNA in respiratory droplets. Our results indicate that surgical face masks could prevent transmission of human coronaviruses and influenza viruses from symptomatic individuals.

Respiratory virus infections cause a broad and overlapping spectrum of symptoms collectively referred to as acute respiratory virus illnesses (ARIs) or more commonly the 'common cold'. Although mostly mild, these ARIs can sometimes cause severe disease and death[1]. These viruses spread between humans through direct or indirect contact, respiratory droplets (including larger droplets that fall rapidly near the source as well as coarse aerosols with aerodynamic diameter >5 μm) and fine-particle aerosols (droplets and droplet nuclei with aerodynamic diameter ≤5 μm)[2,3]. Although hand hygiene and use of face masks, primarily targeting contact and respiratory droplet transmission, have been suggested as important mitigation strategies against influenza virus transmission[4], little is known about the relative importance of these modes in the transmission of other common respiratory viruses[2,3,5]. Uncertainties similarly apply to the modes of transmission of COVID-19 (refs. [6,7]).

Some health authorities recommend that masks be worn by ill individuals to prevent onward transmission (source control)[4,8]. Surgical face masks were originally introduced to protect patients from wound infection and contamination from surgeons (the wearer) during surgical procedures, and were later adopted to protect healthcare workers against acquiring infection from their patients. However, most of the existing evidence on the filtering efficacy of face masks and respirators comes from in vitro experiments with nonbiological particles[9,10], which may not be generalizable to infectious respiratory virus droplets. There is little information on the efficacy of face masks in filtering respiratory viruses and reducing viral release from an individual with respiratory infections[8], and most research has focused on influenza[11,12].

Here we aimed to explore the importance of respiratory droplet and aerosol routes of transmission with a particular focus on coronaviruses, influenza viruses and rhinoviruses, by quantifying the amount of respiratory virus in exhaled breath of participants with medically attended ARIs and determining the potential efficacy of surgical face masks to prevent respiratory virus transmission.

Results

We screened 3,363 individuals in two study phases, ultimately enrolling 246 individuals who provided exhaled breath samples (Extended Data Fig. 1). Among these 246 participants, 122 (50%) participants were randomized to not wearing a face mask during the first exhaled breath collection and 124 (50%) participants were randomized to wearing a face mask. Overall, 49 (20%) voluntarily provided a second exhaled breath collection of the alternate type.

Infections by at least one respiratory virus were confirmed by reverse transcription PCR (RT–PCR) in 123 of 246 (50%) participants. Of these 123 participants, 111 (90%) were infected by human (seasonal) coronavirus ($n=17$), influenza virus ($n=43$) or rhinovirus ($n=54$) (Extended Data Figs. 1 and 2), including one participant co-infected by both coronavirus and influenza virus and another two participants co-infected by both rhinovirus and influenza virus. These 111 participants were the focus of our analyses.

There were some minor differences in characteristics of the 111 participants with the different viruses (Table 1a). Overall, 24% of participants had a measured fever ≥37.8 °C, with patients with influenza more than twice as likely than patients infected with coronavirus and rhinovirus to have a measured fever. Coronavirus-infected participants coughed the most with an average of 17 (s.d. = 30) coughs during the 30-min exhaled breath collection. The profiles of the participants randomized to with-mask versus without-mask groups were similar (Supplementary Table 1).

We tested viral shedding (in terms of viral copies per sample) in nasal swabs, throat swabs, respiratory droplet samples and aerosol samples and compared the latter two between samples collected with or without a face mask (Fig. 1). On average, viral shedding was higher in nasal swabs than in throat swabs for each of coronavirus (median 8.1 $\log_{10}$ virus copies per sample versus 3.9), influenza virus (6.7 versus 4.0) and rhinovirus (6.8 versus 3.3), respectively. Viral RNA was identified from respiratory droplets and aerosols for all three viruses, including 30%, 26% and 28% of respiratory droplets and 40%, 35% and 56% of aerosols collected while not wearing a face mask, from coronavirus, influenza virus and rhinovirus-infected participants, respectively (Table 1b). In particular for coronavirus, we identified OC43 and HKU1 from both respiratory

[1]WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China. [2]Department of Microbiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China. [3]Department of Environmental Health, Harvard School of Public Health, Boston, MA, USA. [4]Department of Surgery, Queen Mary Hospital, Hong Kong, China. [5]Department of Mechanical Engineering, The University of Hong Kong, Hong Kong, China. [6]Department of Pathology, Hong Kong Baptist Hospital, Hong Kong, China. [7]Maryland Institute for Applied Environmental Health, School of Public Health, University of Maryland, College Park, MD, USA. [8]These authors jointly supervised this work: Donald K. Milton, Benjamin J. Cowling. ✉e-mail: bcowling@hku.hk

Table 1a | Characteristics of individuals with symptomatic coronavirus, influenza virus or rhinovirus infection

	All who provided exhaled breath	Coronavirus	Influenza virus	Rhinovirus
	(*n* = 246)	(*n* = 17)	(*n* = 43)	(*n* = 54)
	n (%)	*n* (%)	*n* (%)	*n* (%)
Female	144 (59)	13 (76)	22 (51)	30 (56)
Age group, years				
11–17	12 (5)	0 (0)	8 (19)	4 (7)
18–34	114 (46)	10 (59)	11 (26)	24 (44)
35–50	79 (32)	2 (12)	16 (37)	18 (33)
51–64	35 (14)	4 (24)	8 (19)	5 (9)
≥ 65	6 (2)	1 (6)	0 (0)	3 (6)
Chronic medical conditions				
Any	49 (20)	5 (29)	5 (12)	10 (19)
Respiratory	18 (7)	0 (0)	4 (9)	3 (6)
Influenza vaccination				
Ever	94 (38)	6 (35)	15 (35)	20 (37)
Current season	23 (9)	2 (12)	1 (2)	4 (7)
Previous season only	71 (29)	4 (24)	14 (33)	16 (30)
Ever smoker	31 (13)	1 (6)	6 (14)	6 (11)
Time since illness onset, h				
<24	22 (9)	0 (0)	5 (12)	2 (4)
24–48	100 (41)	9 (53)	13 (30)	25 (46)
48–72	85 (35)	8 (47)	18 (42)	20 (37)
72–96	39 (16)	0 (0)	7 (16)	7 (13)
History of measured fever ≥37.8 °C	58 (24)	3 (18)	17 (40)	8 (15)
Measured fever ≥37.8 °C at presentation	36 (15)	2 (12)	18 (42)	2 (4)
Measured body temperature (˚C) at enrollment (mean, s.d.)	36.8 (0.8)	36.9 (0.8)	37.4 (0.9)	36.6 (0.7)
Symptoms at presentation				
Fever	111 (45)	10 (59)	27 (63)	16 (30)
Cough	198 (80)	15 (88)	40 (93)	44 (81)
Sore throat	211 (86)	15 (88)	31 (72)	49 (91)
Runny nose	200 (81)	17 (100)	36 (84)	48 (89)
Headache	186 (76)	13 (76)	30 (70)	38 (70)
Myalgia	176 (72)	12 (71)	31 (72)	34 (63)
Phlegm	176 (72)	9 (53)	34 (79)	41 (76)
Chest tightness	64 (26)	3 (18)	12 (28)	9 (17)
Shortness of breath	103 (42)	6 (35)	14 (33)	25 (46)
Chills	100 (41)	8 (47)	29 (67)	16 (30)
Sweating	95 (39)	5 (29)	18 (42)	20 (37)
Fatigue	218 (89)	16 (94)	38 (88)	48 (89)
Vomiting	19 (8)	2 (12)	5 (12)	2 (4)
Diarrhea	17 (7)	2 (12)	1 (2)	6 (11)
Number of coughs during exhaled breath collection (mean, s.d.)	8 (14)	17 (30)	8 (11)	5 (9)

Seasonal coronavirus (*n* = 17), seasonal influenza virus (*n* = 43) and rhinovirus (*n* = 54) infections were confirmed in individuals with acute respiratory symptoms by RT-PCR in any samples (nasal swab, throat swab, respiratory droplets and aerosols) collected.

droplets and aerosols, but only identified NL63 from aerosols and not from respiratory droplets (Supplementary Table 2 and Extended Data Fig. 3).

We detected coronavirus in respiratory droplets and aerosols in 3 of 10 (30%) and 4 of 10 (40%) of the samples collected without face masks, respectively, but did not detect any virus in respiratory droplets or aerosols collected from participants wearing face masks, this difference was significant in aerosols and showed a trend toward reduced detection in respiratory droplets (Table 1b). For influenza virus, we detected virus in 6 of 23 (26%) and 8 of 23 (35%) of the

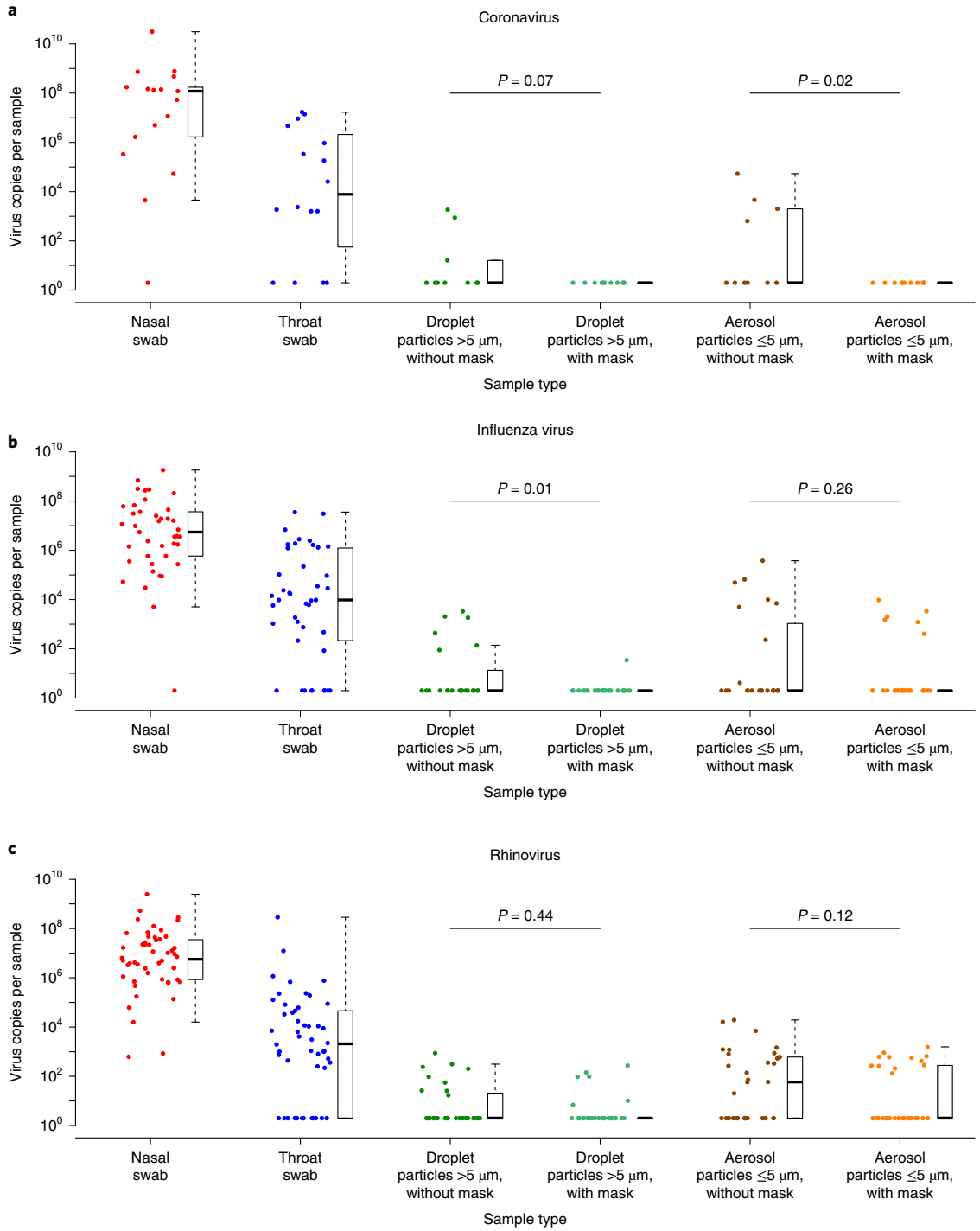


Fig. 1 | Efficacy of surgical face masks in reducing respiratory virus shedding in respiratory droplets and aerosols of symptomatic individuals with coronavirus, influenza virus or rhinovirus infection. a–**c**, Virus copies per sample collected in nasal swab (red), throat swab (blue) and respiratory droplets collected for 30min while not wearing (dark green) or wearing (light green) a surgical face mask, and aerosols collected for 30min while not wearing (brown) or wearing (orange) a face mask, collected from individuals with acute respiratory symptoms who were positive for coronavirus (**a**), influenza virus (**b**) and rhinovirus (**c**), as determined by RT–PCR in any samples. *P* values for mask intervention as predictor of $\log_{10}$ virus copies per sample in an unadjusted univariate Tobit regression model which allowed for censoring at the lower limit of detection of the RT–PCR assay are shown, with significant differences in bold. For nasal swabs and throat swabs, all infected individuals were included (coronavirus, *n*=17; influenza virus, *n*=43; rhinovirus, *n*=54). For respiratory droplets and aerosols, numbers of infected individuals who provided exhaled breath samples while not wearing or wearing a surgical face mask, respectively were: coronavirus (*n*=10 and 11), influenza virus (*n*=23 and 28) and rhinovirus (*n*=36 and 32). A subset of participants provided exhaled breath samples for both mask interventions (coronavirus, *n*=4; influenza virus, *n*=8; rhinovirus, *n*=14). The box plots indicate the median with the interquartile range (lower and upper hinge) and ±1.5×interquartile range from the first and third quartile (lower and upper whiskers).

Table 1b | Efficacy of surgical face masks in reducing respiratory virus frequency of detection and viral shedding in respiratory droplets and aerosols of symptomatic individuals with coronavirus, influenza virus or rhinovirus infection

	Droplet particles >5 µm			Aerosol particles ≤5 µm		
Virus type	**Without surgical face mask**	**With surgical face mask**	***P***	**Without surgical face mask**	**With surgical face mask**	***P***
	Detection of virus					
	No. positive/no. total (%)	**No. positive/no. total (%)**		**No. positive/no. total (%)**	**No. positive/no. total (%)**	
Coronavirus	3 of 10 (30)	0 of 11 (0)	0.09	**4 of 10 (40)**	**0 of 11 (0)**	**0.04**
Influenza virus	**6 of 23 (26)**	**1 of 27 (4)**	**0.04**	8 of 23 (35)	6 of 27 (22)	0.36
Rhinovirus	9 of 32 (28)	6 of 27 (22)	0.77	19 of 34 (56)	12 of 32 (38)	0.15
	Viral load (log_{10} virus copies per sample)					
	Median (IQR)	**Median (IQR)**		**Median (IQR)**	**Median (IQR)**	
Coronavirus	0.3 (0.3, 1.2)	0.3 (0.3, 0.3)	0.07	**0.3 (0.3, 3.3)**	**0.3 (0.3, 0.3)**	**0.02**
Influenza virus	**0.3 (0.3, 1.1)**	**0.3 (0.3, 0.3)**	**0.01**	0.3 (0.3, 3.0)	0.3 (0.3, 0.3)	0.26
Rhinovirus	0.3 (0.3, 1.3)	0.3 (0.3, 0.3)	0.44	1.8 (0.3, 2.8)	0.3 (0.3, 2.4)	0.12

P values for comparing the frequency of respiratory virus detection between the mask intervention were obtained by two-sided Fisher's exact test and (two-sided) *P* values for mask intervention as predictor of log_{10} virus copies per sample were obtained by an unadjusted univariate Tobit regression model, which allowed for censoring at the lower limit of detection of the RT-PCR assay, with significant differences in bold. Undetectable values were imputed as 0.3 log_{10} virus copies per sample. IQR, interquartile range.

respiratory droplet and aerosol samples collected without face masks, respectively. There was a significant reduction by wearing face masks to 1 of 27 (4%) in detection of influenza virus in respiratory droplets, but no significant reduction in detection in aerosols (Table 1b). Moreover, among the eight participants who had influenza virus detected by RT–PCR from without-mask aerosols, five were tested by viral culture and four were culture-positive. Among the six participants who had influenza virus detected by RT–PCR from with-mask aerosols, four were tested by viral culture and two were culture-positive. For rhinovirus, there were no significant differences between detection of virus with or without face masks, both in respiratory droplets and in aerosols (Table 1b). Conclusions were similar in comparisons of viral shedding (Table 1b). In addition, we found a significant reduction in viral shedding (Supplementary Table 2) in respiratory droplets for OC43 (Extended Data Fig. 4) and influenza B virus (Extended Data Fig. 5) and in aerosols for NL63 (Extended Data Fig. 4).

We identified correlations between viral loads in different samples (Extended Data Figs. 6–8) and some evidence of declines in viral shedding by time since onset for influenza virus but not for coronavirus or rhinovirus (Extended Data Fig. 9). In univariable analyses of factors associated with detection of respiratory viruses in various sample types, we did not identify significant association in viral shedding with days since symptom onset (Supplementary Table 3) for respiratory droplets or aerosols (Supplementary Tables 4–6).

A subset of participants (72 of 246, 29%) did not cough at all during at least one exhaled breath collection, including 37 of 147 (25%) during the without-mask and 42 of 148 (28%) during the with-mask breath collection. In the subset for coronavirus ($n=4$), we did not detect any virus in respiratory droplets or aerosols from any participants. In the subset for influenza virus ($n=9$), we detected virus in aerosols but not respiratory droplets from one participant. In the subset for rhinovirus ($n=17$), we detected virus in respiratory droplets from three participants, and we detected virus in aerosols in five participants.

Discussion

Our results indicate that aerosol transmission is a potential mode of transmission for coronaviruses as well as influenza viruses and rhinoviruses. Published studies detected respiratory viruses[13,14] such as influenza[12,15] and rhinovirus[16] from exhaled breath, and the detection of SARS-CoV[17] and MERS-CoV[18] from air samples (without size fractionation) collected from hospitals treating patients with severe acute respiratory syndrome and Middle East respiratory syndrome, but ours demonstrates detection of human seasonal coronaviruses in exhaled breath, including the detection of OC43 and HKU1 from respiratory droplets and NL63, OC43 and HKU1 from aerosols.

Our findings indicate that surgical masks can efficaciously reduce the emission of influenza virus particles into the environment in respiratory droplets, but not in aerosols[12]. Both the previous and current study used a bioaerosol collecting device, the Gesundheit-II (G-II)[12,15,19], to capture exhaled breath particles and differentiated them into two size fractions, where exhaled breath coarse particles >5 µm (respiratory droplets) were collected by impaction with a 5-µm slit inertial Teflon impactor and the remaining fine particles ≤5 µm (aerosols) were collected by condensation in buffer. We also demonstrated the efficacy of surgical masks to reduce coronavirus detection and viral copies in large respiratory droplets and in aerosols (Table 1b). This has important implications for control of COVID-19, suggesting that surgical face masks could be used by ill people to reduce onward transmission.

Among the samples collected without a face mask, we found that the majority of participants with influenza virus and coronavirus infection did not shed detectable virus in respiratory droplets or aerosols, whereas for rhinovirus we detected virus in aerosols in 19 of 34 (56%) participants (compared to 4 of 10 (40%) for coronavirus and 8 of 23 (35%) for influenza). For those who did shed virus in respiratory droplets and aerosols, viral load in both tended to be low (Fig. 1). Given the high collection efficiency of the G-II (ref. [19]) and given that each exhaled breath collection was conducted for 30 min, this might imply that prolonged close contact would be required for transmission to occur, even if transmission was primarily via aerosols, as has been described for rhinovirus colds[20]. Our results also indicate that there could be considerable heterogeneity in contagiousness of individuals with coronavirus and influenza virus infections.

The major limitation of our study was the large proportion of participants with undetectable viral shedding in exhaled breath for each of the viruses studied. We could have increased the sampling duration beyond 30 min to increase the viral shedding being captured, at the cost of acceptability in some participants. An alternative approach would be to invite participants to perform forced coughs during exhaled breath collection[12]. However, it was the aim of our present study to focus on recovering respiratory

virus in exhaled breath in a real-life situation and we expected that some individuals during an acute respiratory illness would not cough much or at all. Indeed, we identified virus RNA in a small number of participants who did not cough at all during the 30-min exhaled breath collection, which would suggest droplet and aerosol routes of transmission are possible from individuals with no obvious signs or symptoms. Another limitation is that we did not confirm the infectivity of coronavirus or rhinovirus detected in exhaled breath. While the G-II was designed to preserve viability of viruses in aerosols, and in the present study we were able to identify infectious influenza virus in aerosols, we did not attempt to culture coronavirus or rhinovirus from the corresponding aerosol samples.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-020-0843-2.

Received: 2 March 2020; Accepted: 20 March 2020;
Published online: 3 April 2020

References

1. Nichols, W. G., Peck Campbell, A. J. & Boeckh, M. Respiratory viruses other than influenza virus: impact and therapeutic advances. *Clin. Microbiol. Rev.* **21**, 274–290 (2008).
2. Shiu, E. Y. C., Leung, N. H. L. & Cowling, B. J. Controversy around airborne versus droplet transmission of respiratory viruses: implication for infection prevention. *Curr. Opin. Infect. Dis.* **32**, 372–379 (2019).
3. Tellier, R., Li, Y., Cowling, B. J. & Tang, J. W. Recognition of aerosol transmission of infectious agents: a commentary. *BMC Infect. Dis.* **19**, 101 (2019).
4. Xiao, J. et al. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings-personal protective and environmental measures. *Emerg. Infect. Dis.* https://doi.org/10.3201/eid2605.190994 (2020).
5. Kutter, J. S., Spronken, M. I., Fraaij, P. L., Fouchier, R. A. M. & Herfst, S. Transmission routes of respiratory viruses among humans. *Curr. Opin. Virol.* **28**, 142–151 (2018).
6. Cowling, B. J. & Leung, G. M. Epidemiological research priorities for public health control of the ongoing global novel coronavirus (2019-nCoV) outbreak. *Euro Surveill.* https://doi.org/10.2807/1560-7917.ES.2020.25.6.2000110 (2020).
7. Han, Q., Lin, Q., Ni, Z. & You, L. Uncertainties about the transmission routes of 2019 novel coronavirus. *Influenza Other Respir. Viruses* https://doi.org/10.1111/irv.12735 (2020).
8. MacIntyre, C. R. & Chughtai, A. A. Facemasks for the prevention of infection in healthcare and community settings. *BMJ* **350**, h694 (2015).
9. Ha'eri, G. B. & Wiley, A. M. The efficacy of standard surgical face masks: an investigation using "tracer particles". *Clin. Orthop. Relat. Res.* **148**, 160–162 (1980).
10. Patel, R. B., Skaria, S. D., Mansour, M. M. & Smaldone, G. C. Respiratory source control using a surgical mask: an in vitro study. *J. Occup. Environ. Hyg.* **13**, 569–576 (2016).
11. Johnson, D. F., Druce, J. D., Birch, C. & Grayson, M. L. A quantitative assessment of the efficacy of surgical and N95 masks to filter influenza virus in patients with acute influenza infection. *Clin. Infect. Dis.* **49**, 275–277 (2009).
12. Milton, D. K., Fabian, M. P., Cowling, B. J., Grantham, M. L. & McDevitt, J. J. Influenza virus aerosols in human exhaled breath: particle size, culturability, and effect of surgical masks. *PLoS Pathog.* **9**, e1003205 (2013).
13. Huynh, K. N., Oliver, B. G., Stelzer, S., Rawlinson, W. D. & Tovey, E. R. A new method for sampling and detection of exhaled respiratory virus aerosols. *Clin. Infect. Dis.* **46**, 93–95 (2008).
14. Stelzer-Braid, S. et al. Exhalation of respiratory viruses by breathing, coughing and talking. *J. Med. Virol.* **81**, 1674–1679 (2009).
15. Yan, J. et al. Infectious virus in exhaled breath of symptomatic seasonal influenza cases from a college community. *Proc. Natl Acad. Sci. USA* **115**, 1081–1086 (2018).
16. Tovey, E. R. et al. Rhinoviruses significantly affect day-to-day respiratory symptoms of children with asthma. *J. Allergy Clin. Immunol.* **135**, 663–669 (2015).
17. Booth, T. F. et al. Detection of airborne severe acute respiratory syndrome (SARS) coronavirus and environmental contamination in SARS outbreak units. *J. Infect. Dis.* **191**, 1472–1477 (2005).
18. Kim, S. H. et al. Extensive viable Middle East respiratory syndrome (MERS) coronavirus contamination in air and surrounding environment in MERS isolation wards. *Clin. Infect. Dis.* **63**, 363–369 (2016).
19. McDevitt, J. J. et al. Development and performance evaluation of an exhaled-breath bioaerosol collector for influenza virus. *Aerosol Sci. Technol.* **47**, 444–451 (2013).
20. Jennings, L. C. & Dick, E. C. Transmission and control of rhinovirus colds. *Eur. J. Epidemiol.* **3**, 327–335 (1987).

Methods

Study design. Participants were recruited year-round from March 2013 through May 2016 in a general outpatient clinic of a private hospital in Hong Kong. As routine practice, clinic staff screened all individuals attending the clinics for respiratory and any other symptoms regardless of the purpose of the visit at triage. Study staff then approached immediately those who reported at least one of the following symptoms of ARI for further screening: fever ≥37.8 °C, cough, sore throat, runny nose, headache, myalgia and phlegm. Individuals who reported ≥2 ARI symptoms, within 3 d of illness onset and ≥11 years of age were eligible to participate. After explaining the study to and obtaining informed consent from the participants, a rapid influenza diagnostic test, the Sofia Influenza A + B Fluorescent Immunoassay Analyzer (cat. no. 20218, Quidel), was used to identify influenza A or B virus infection as an incentive to participate. All participants provided a nasal swab for the rapid test and an additional nasal swab and a separate throat swab for subsequent virologic confirmation at the laboratory. All participants also completed a questionnaire to record basic information including age, sex, symptom severity, medication, medical conditions and smoking history. In the first phase of the study from March 2013 to February 2014 ('Influenza Study'), the result of the rapid test was used to determine eligibility for further participation in the study and exhaled breath collection, whereas in the second phase of the study from March 2014 to May 2016 ('Respiratory Virus Study'), the rapid test did not affect eligibility. Eligible participants were then invited to provide an exhaled breath sample for 30 min in the same clinic visit.

Before exhaled breath collection, each participant was randomly allocated in a 1:1 ratio to either wearing a surgical face mask (cat. no. 62356, Kimberly-Clark) or not during the collection. To mimic the real-life situation, under observation by the study staff, participants were asked to attach the surgical mask themselves, but instruction on how to wear the mask properly was given when the participant wore the mask incorrectly. Participants were instructed to breathe as normal during the collection, but (natural) coughing was allowed and the number of coughs was recorded by study staff. Participants were then invited to provide a second exhaled breath sample of the alternate type (for example if the participant was first assigned to wearing a mask they would then provide a second sample without a mask), but most participants did not agree to stay for a second measurement because of time constraints. Participants were compensated for each 30-min exhaled breath collection with a supermarket coupon worth approximately US$30 and all participants were gifted a tympanic thermometer worth approximately US$20.

Ethical approval. Written informed consent was obtained from all participants ≥18 years of age and written informed consent was obtained from parents or legal guardians of participants 11–17 years of age in addition to their own written informed consent. The study protocol was approved by the Institutional Review Board of The University of Hong Kong and the Clinical and Research Ethics Committee of Hong Kong Baptist Hospital.

Collection of swabs and exhaled breath particles. Nasal swabs and throat swabs were collected separately, placed in virus transport medium, stored and transported to the laboratory at 2–8 °C and the virus transport medium was aliquoted and stored at −70 °C until further analysis. Exhaled breath particles were captured and differentiated into two size fractions, the coarse fraction containing particles with aerodynamic diameter >5 µm (referred to here as 'respiratory droplets'), which included droplets up to approximately 100 µm in diameter and the fine fraction with particles ≤5 µm (referred to here as 'aerosols') by the G-II bioaerosol collecting device[12,15,19]. In the G-II device, exhaled breath coarse particles >5 µm were collected by a 5-µm slit inertial Teflon impactor and the remaining fine particles ≤5 µm were condensed and collected into approximately 170 ml of 0.1% BSA/PBS. Both the impactor and the condensate were stored and transported to the laboratory at 2–8 °C. The virus on the impactor was recovered into 1 ml and the condensate was concentrated into 2 ml of 0.1% BSA/PBS, aliquoted and stored at −70 °C until further analysis. In a validation study, the G-II was able to recover over 85% of fine particles >0.05 µm in size and had comparable collection efficiency of influenza virus as the SKC BioSampler[19].

Laboratory testing. Samples collected from the two studies were tested at the same time. Nasal swab samples were first tested by a diagnostic-use viral panel, xTAG Respiratory Viral Panel (Abbott Molecular) to qualitatively detect 12 common respiratory viruses and subtypes including coronaviruses (NL63, OC43, 229E and HKU1), influenza A (nonspecific, H1 and H3) and B viruses, respiratory syncytial virus, parainfluenza virus (types 1–4), adenovirus, human metapneumovirus and enterovirus/rhinovirus. After one or more of the candidate respiratory viruses was detected by the viral panel from the nasal swab, all the samples from the same participant (nasal swab, throat swab, respiratory droplets and aerosols) were then tested with RT–PCR specific for the candidate virus(es) for determination of virus concentration in the samples. Infectious influenza virus was identified by viral culture using MDCK cells as described previously[21], whereas viral culture was not performed for coronavirus and rhinovirus.

Statistical analyses. The primary outcome of the study was virus generation rate in tidal breathing of participants infected by different respiratory viruses and the efficacy of face masks in preventing virus dissemination in exhaled breath, separately considering the respiratory droplets and aerosols. The secondary outcomes were correlation between viral shedding in nose swabs, throat swabs, respiratory droplets and aerosols and factors affecting viral shedding in respiratory droplets and aerosols.

We identified three groups of respiratory viruses with the highest frequency of infection as identified by RT–PCR, namely coronavirus (including NL63, OC43, HKU1 and 229E), influenza virus and rhinovirus, for further statistical analyses. We defined viral shedding as $\log_{10}$ virus copies per sample and plotted viral shedding in each sample (nasal swab, throat swab, respiratory droplets and aerosols); the latter two were stratified by mask intervention. As a proxy for the efficacy of face masks in preventing transmission of respiratory viruses via respiratory droplet and aerosol routes, we compared the respiratory virus viral shedding in respiratory droplet and aerosol samples between participants wearing face masks or not, by comparing the frequency of detection with a two-sided Fisher's exact test and by comparing viral load (defined as $\log_{10}$ virus copies per sample) by an unadjusted univariate Tobit regression model, which allowed for censoring at the lower limit of detection of the RT–PCR assay. We also used the unadjusted univariate Tobit regression to investigate factors affecting viral shedding in respiratory droplets and aerosols without mask use, for example age, days since symptom onset, previous influenza vaccination, current medication and number of coughs during exhaled breath collection. We investigated correlations between viral shedding in nasal swab, throat swab, respiratory droplets and aerosols with scatter-plots and calculated the Spearman's rank correlation coefficient between any two types of samples. We imputed 0.3 $\log_{10}$ virus copies ml^{-1} for undetectable values before transformation to $\log_{10}$ virus copies per sample. All analyses were conducted with R v.3.6.0 (ref. [22]) and the VGAM package v.1.1.1 (ref. [23]).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Anonymized raw data and R syntax to reproduce all the analyses, figures, tables and supplementary tables in the published article are available at: https://doi.org/10.5061/dryad.w9ghx3fkt.

References

21. Chan, K. H., Peiris, J. S., Lim, W., Nicholls, J. M. & Chiu, S. S. Comparison of nasopharyngeal flocked swabs and aspirates for rapid diagnosis of respiratory viruses in children. *J. Clin. Virol.* **42**, 65–69 (2008).
22. R: a language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, 2019).
23. Yee, T. W. *Vector Generalized Linear and Additive Models: with an Implementation in R* (Springer, 2016).

Acknowledgements

This work was supported by the General Research Fund of the University Grants Committee (grant no. 765811), the Health and Medical Research Fund (grant no. 13120592) and a commissioned grant of the Food and Health Bureau and the Theme-based Research Scheme (project no. T11-705/14-N) of the Research Grants Council of the Hong Kong SAR Government. We acknowledge colleagues including R. O. P. Fung, A. K. W. Li, T. W. Y. Ng, T. H. C. So, P. Wu and Y. Xie for technical support in preparing and conducting this study and enrolling participants; J. K. M. Chan, S. Y. Ho, Y. Z. Liu and A. Yu for laboratory support; S. Ferguson, W. K. Leung, J. Pantelic, J. Wei and M. Wolfson for technical support in constructing and maintaining the G-II device; V. J. Fang, L. M. Ho and T. T. K. Lui for setting up the database; and C. W. Y. Cheung, L. F. K. Cheung, P. T. Y. Ching, A. C. H. Lai, D. W. Y. Lam, S. S. Y. Lo, A. S. K. Luk and other colleagues at the Outpatient Center and Infection Control Team of Hong Kong Baptist Hospital for facilitating this study.

Author contributions

All authors meet the International Committee of Medical Journal Editors criteria for authorship. The study protocol was drafted by N.H.L.L. and B.J.C. Data were collected by N.H.L.L., E.Y.C.S. and B.J.P.H. Laboratory testing was performed by D.K.W.C. and K.-H.C. Statistical analyses were conducted by N.H.L.L. N.H.L.L. and B.J.C. wrote the first draft of the manuscript, and all authors provided critical review and revision of the text and approved the final version.

Competing interests

B.J.C. consults for Roche and Sanofi Pasteur. The authors declare no other competing interests.

Additional information

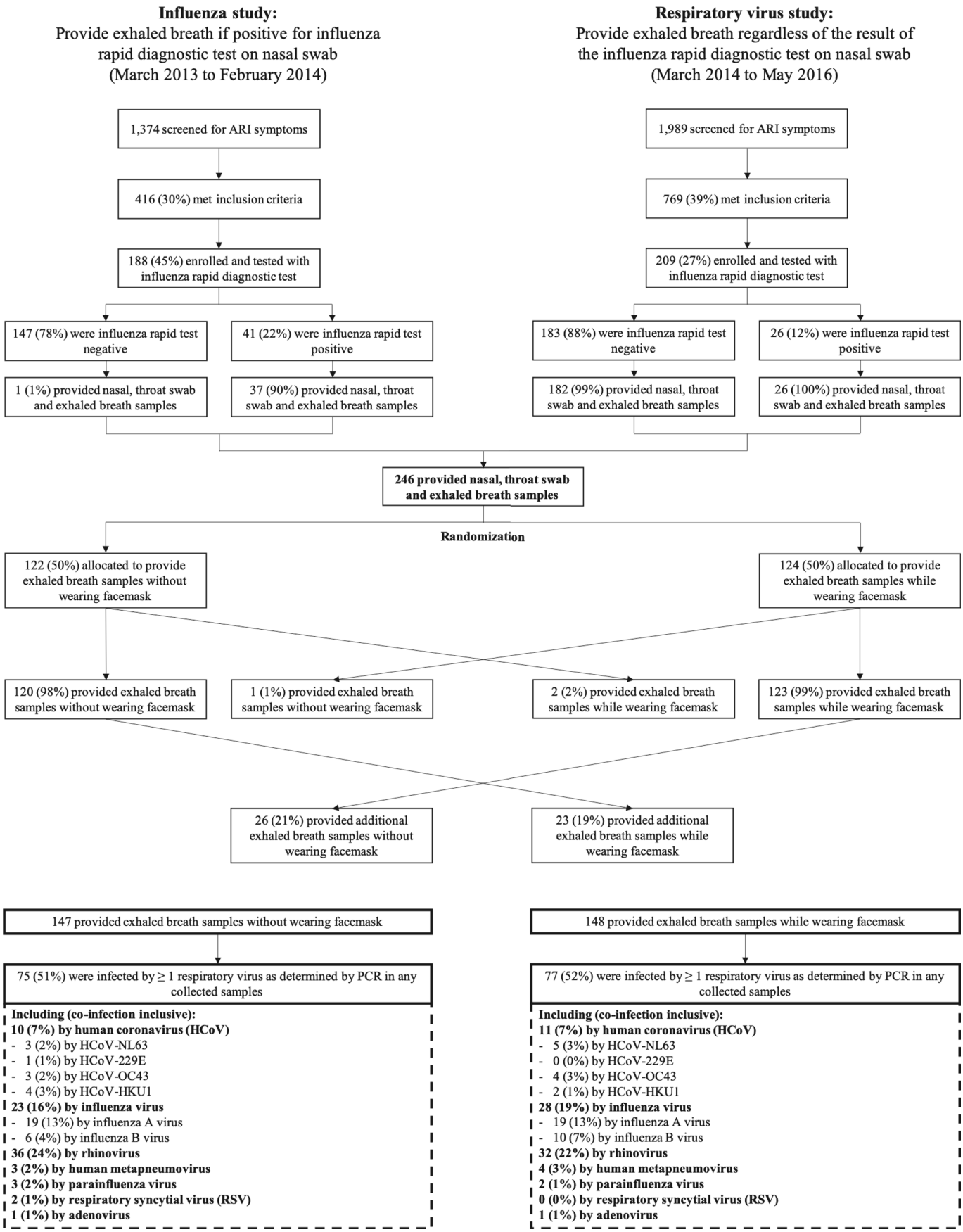
Extended data is available for this paper at https://doi.org/10.1038/s41591-020-0843-2.

Supplementary information is available for this paper at https://doi.org/10.1038/s41591-020-0843-2.

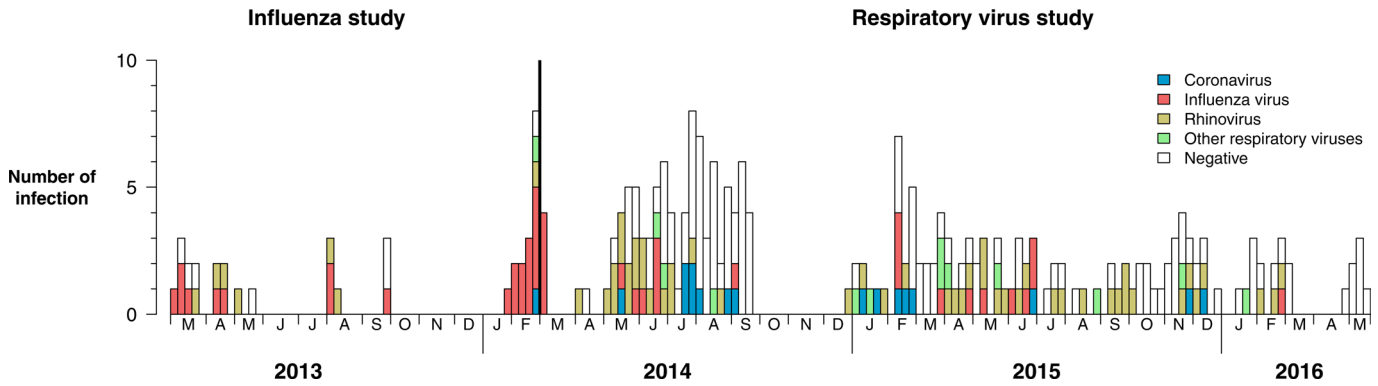
Correspondence and requests for materials should be addressed to B.J.C.

Peer review information Alison Farrell was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

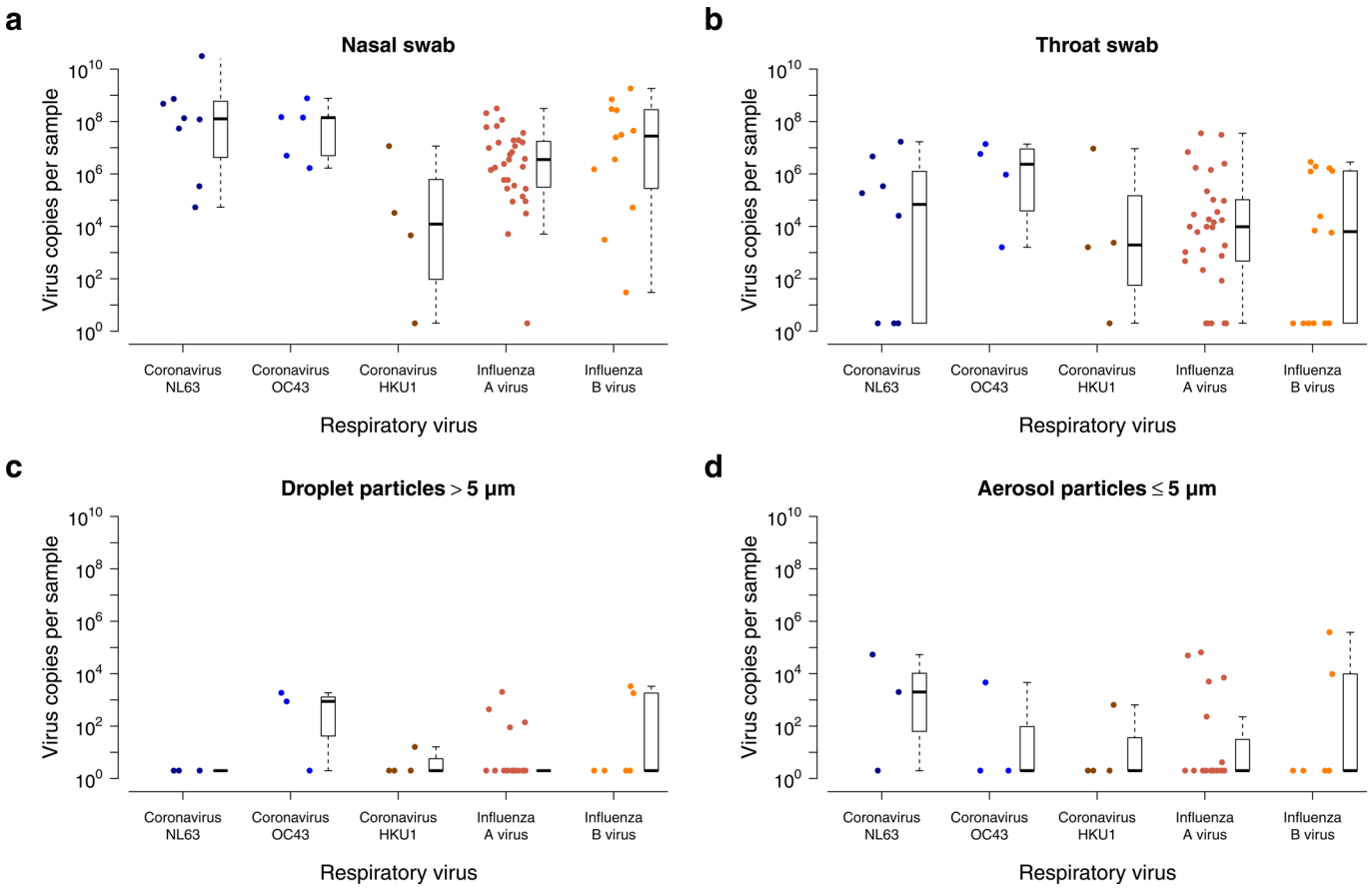
Reprints and permissions information is available at www.nature.com/reprints.



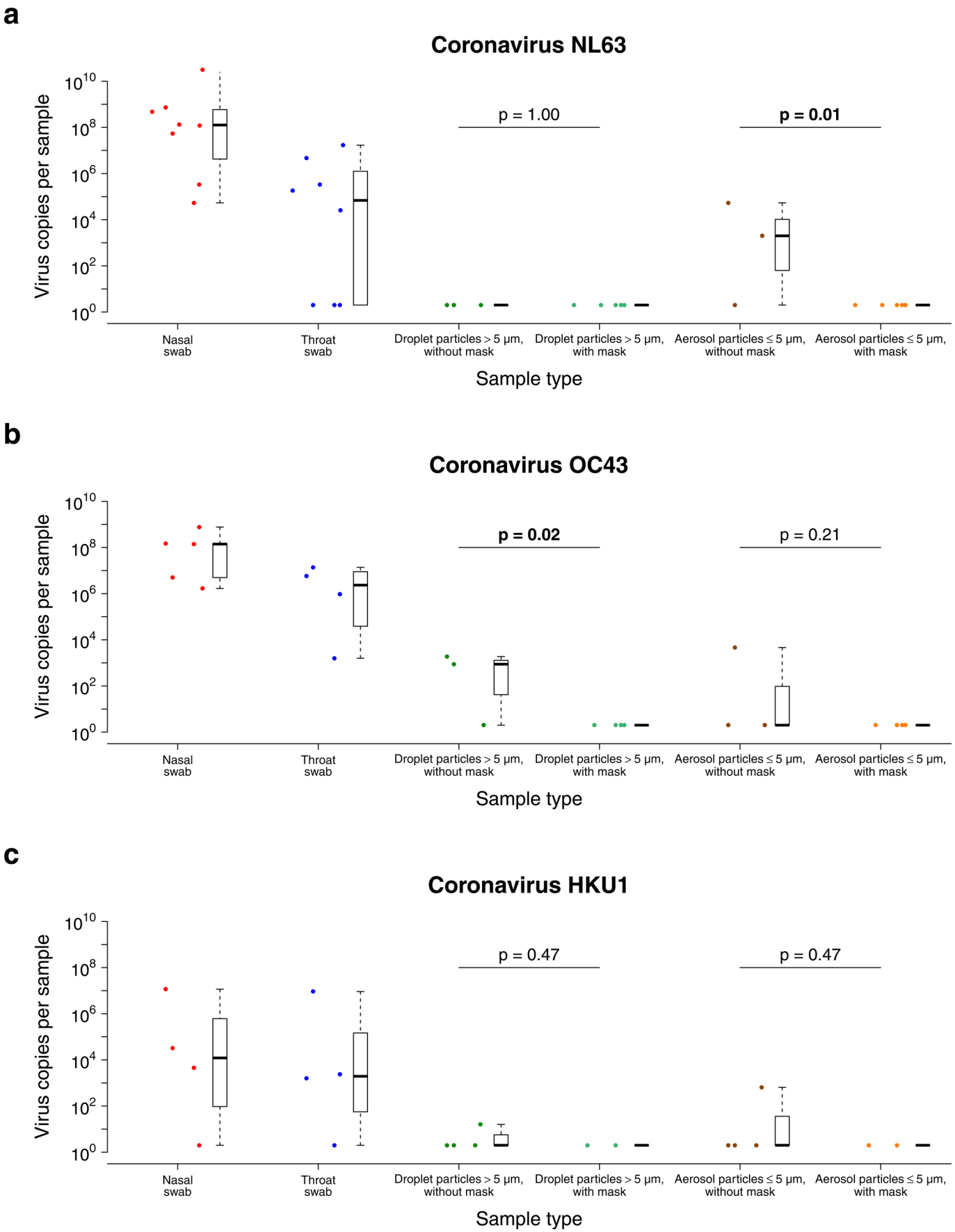
Extended Data Fig. 1 | Participant enrolment, randomization of mask intervention and identification of respiratory virus infection.



Extended Data Fig. 2 | Weekly number of respiratory virus infections identified by RT-PCR in symptomatic individuals who had provided exhaled breath samples (respiratory droplets and aerosols) during the study period. Blue, coronavirus; red, influenza virus; yellow, rhinovirus; green, other respiratory viruses including human metapneumovirus, parainfluenza virus, respiratory syncytial virus and adenovirus; white, no respiratory virus infection identified.

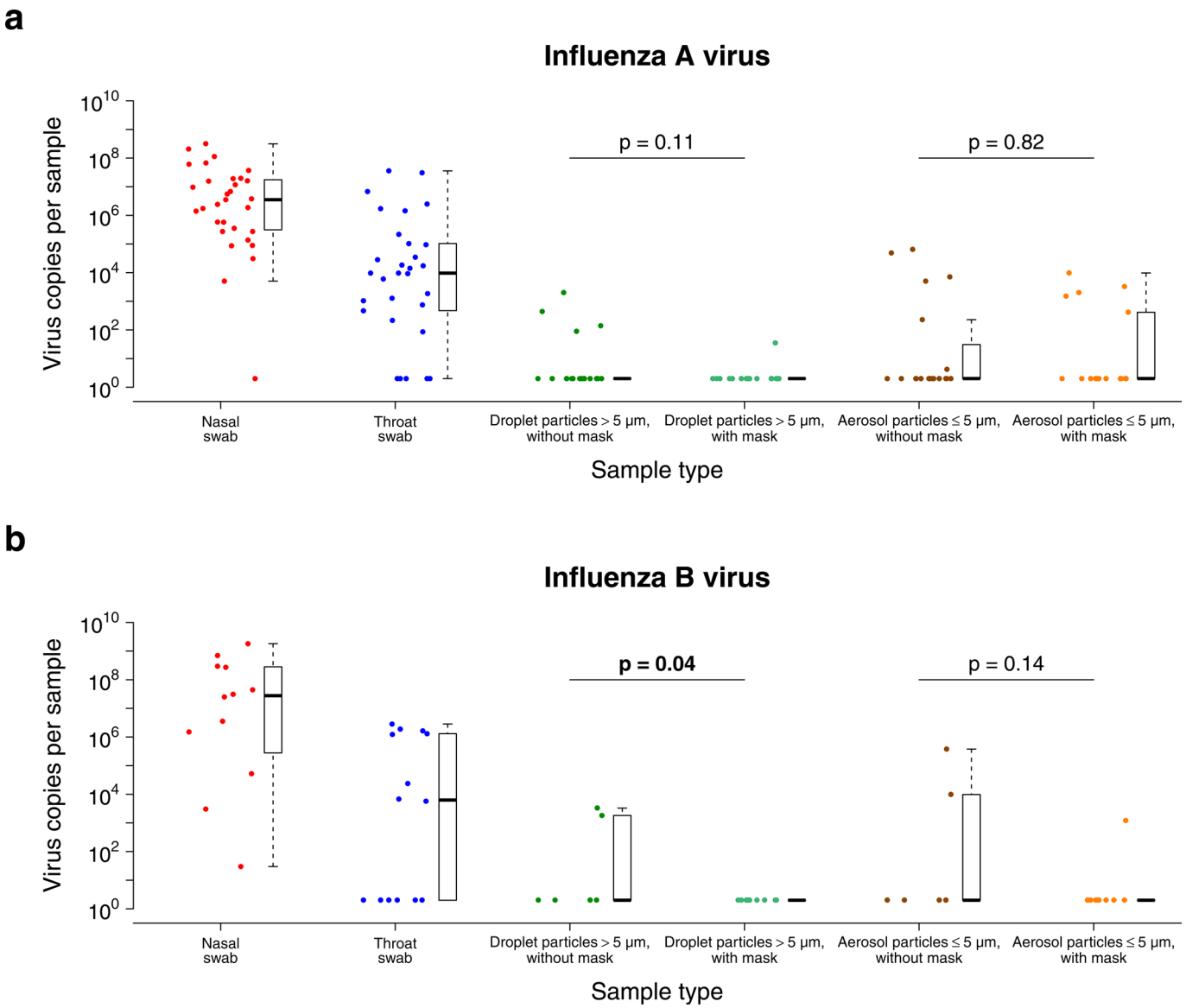


Extended Data Fig. 3 | Respiratory virus shedding in (a) nasal swab, (b) throat swab, (c) respiratory droplets and (d) aerosols in symptomatic individuals with coronavirus NL63, coronavirus OC43, coronavirus HKU1, influenza A and influenza B virus infection. For nasal swabs and throat swabs, all infected individuals identified by RT-PCR in any collected samples were included: coronavirus NL63 (n = 8), coronavirus OC43 (n = 5), coronavirus HKU1 (n = 4), influenza A virus (n = 31) and influenza B virus (n = 14). For respiratory droplets and aerosols, only infected individuals who provided exhaled breath samples while not wearing a surgical face mask were included: coronavirus NL63 (n = 3), coronavirus OC43 (n = 3), coronavirus HKU1 (n = 4), influenza A virus (n = 19) and influenza B virus (n = 6). The box plots indicate the median with the interquartile range (lower and upper hinge) and ±1.5 × interquartile range from the first and third quartile (lower and upper whisker). Dark blue, coronavirus NL63; light blue, coronavirus OC43; brown, coronavirus HKU1; red, influenza A virus; orange, influenza B virus.

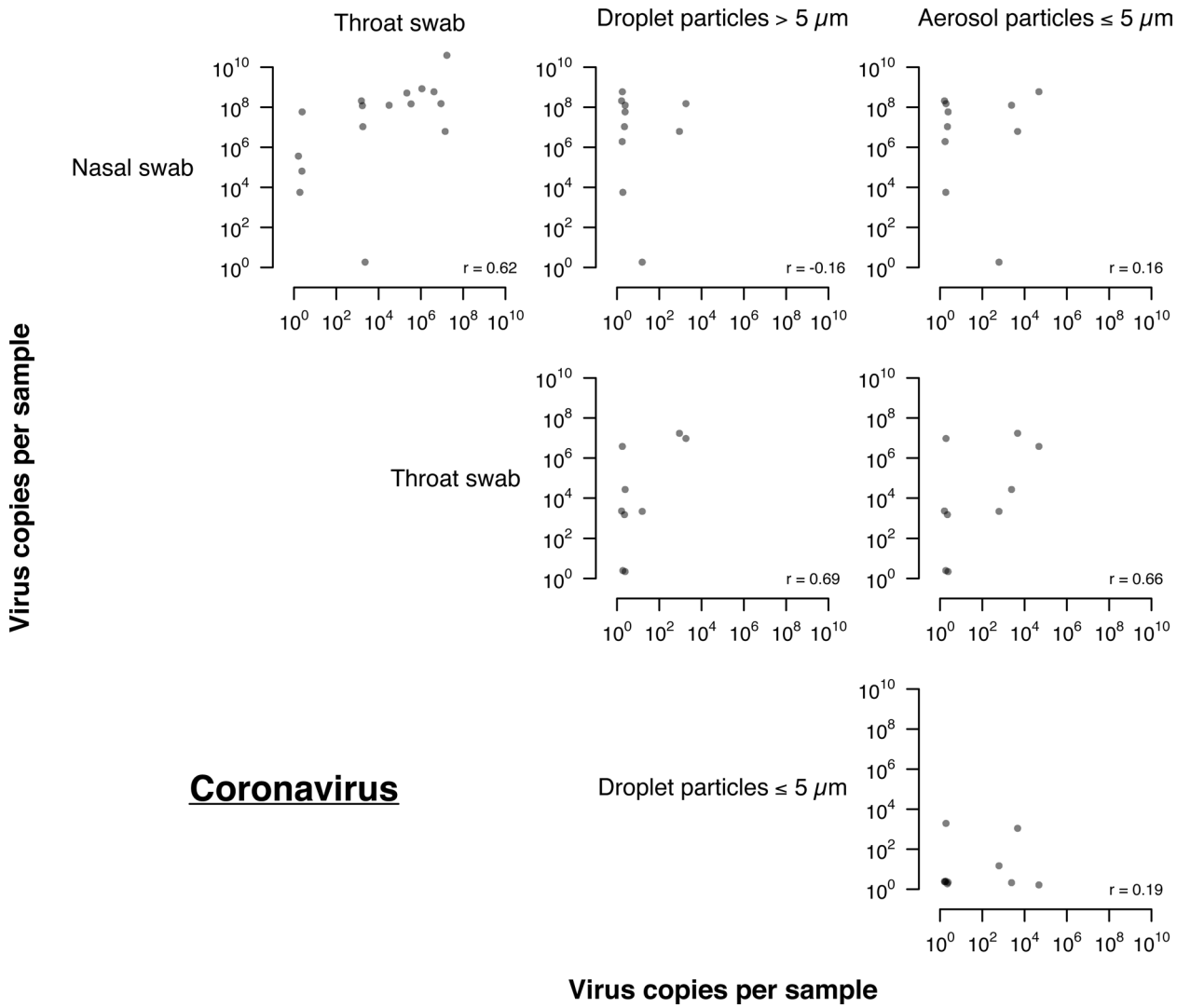


Extended Data Fig. 4 | See next page for caption.

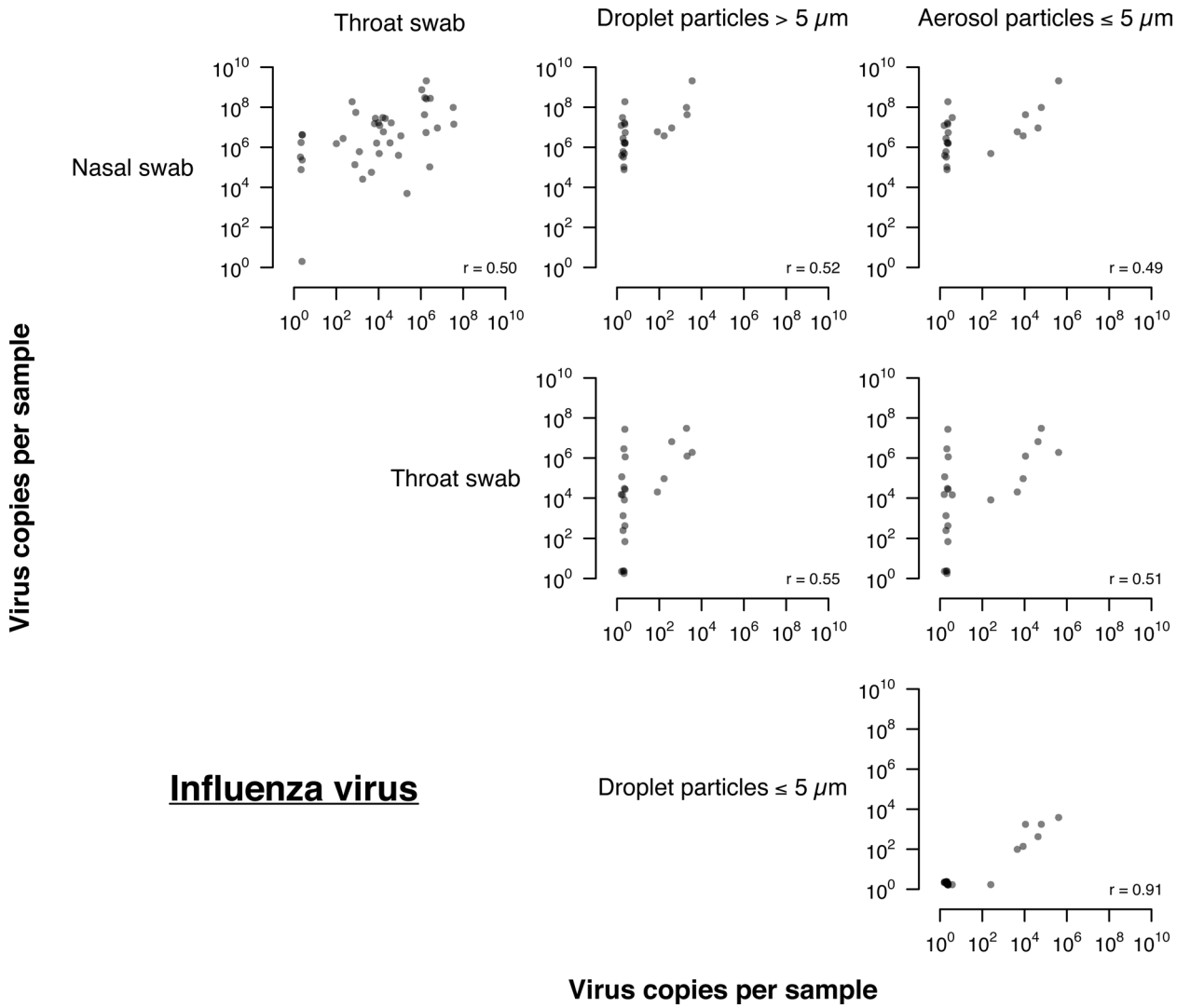
Extended Data Fig. 4 | Efficacy of surgical face masks in reducing respiratory virus shedding in respiratory droplets and aerosols of symptomatic individuals with seasonal coronaviruses including (a) coronavirus NL63, (b) coronavirus OC43 and (c) coronavirus HKU1. The figure shows the virus copies per sample collected in nasal swab (red), throat swab (blue), respiratory droplets collected for 30 min while not wearing (dark green) or wearing (light green) a surgical face mask and aerosols collected for 30 min while not wearing (brown) or wearing (orange) a face mask, collected from individuals with acute respiratory symptoms who were positive for coronavirus NL63, coronavirus OC43 and coronavirus HKU1 as determined by RT-PCR in any samples. *P* values for mask intervention as predictor of $\log_{10}$ virus copies per sample in an unadjusted univariate Tobit regression model which allowed for censoring at the lower limit of detection of the RT-PCR assay are shown, with significant differences in bold. For nasal swabs and throat swabs, all infected individuals were included (coronavirus NL63, n = 8; coronavirus OC43, n = 5; coronavirus HKU1, n = 4). For respiratory droplets and aerosols, numbers of infected individuals who provided exhaled breath samples while not wearing or wearing a surgical face mask, respectively were: coronavirus NL63 (n = 3 and 5), coronavirus OC43 (n = 3 and 4), coronavirus HKU1 (n = 4 and 2). A subset of participants provided exhaled breath samples for both mask interventions (coronavirus NL63, n = 0; coronavirus OC43, n = 2; coronavirus HKU1, n = 2).



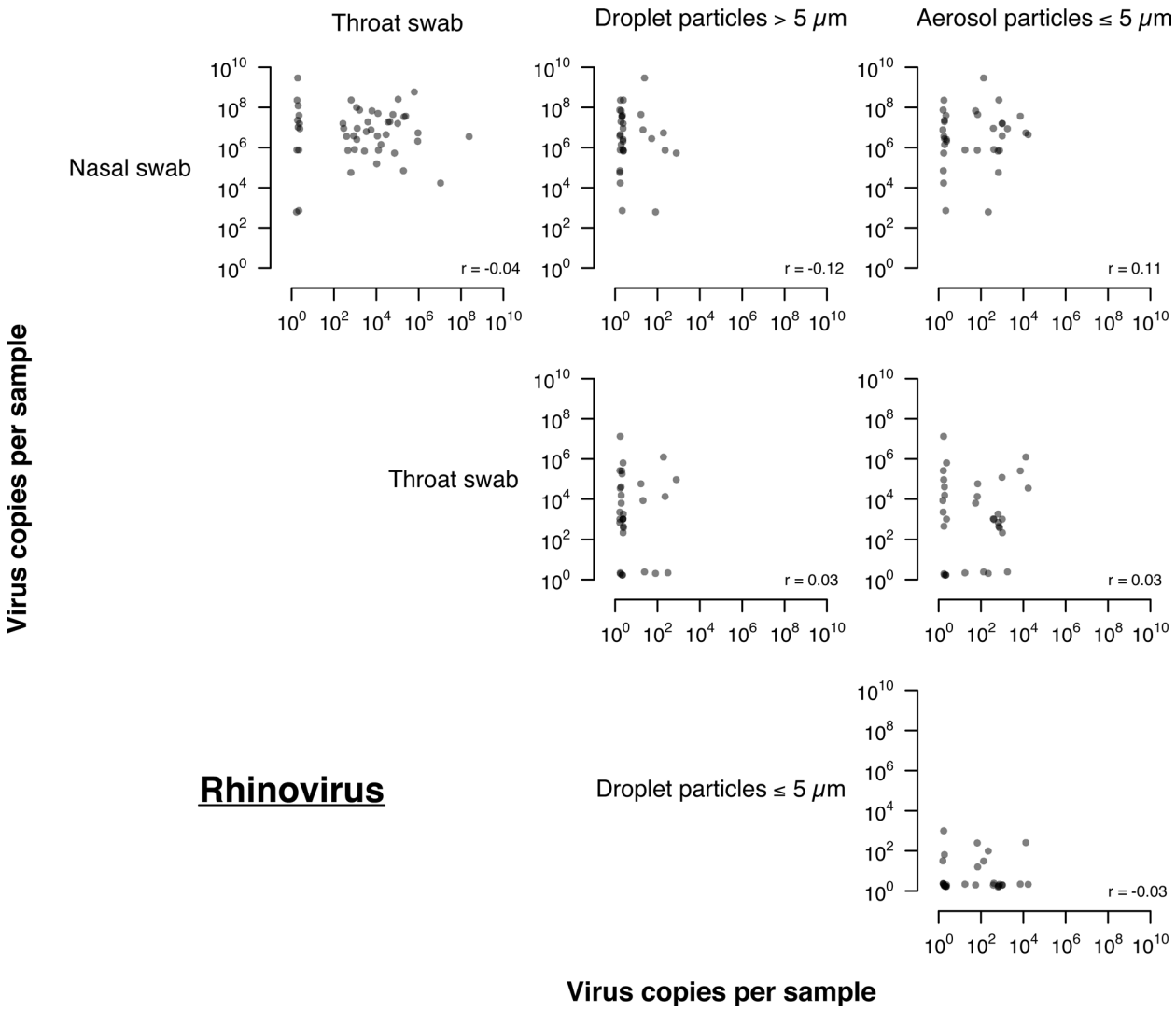
Extended Data Fig. 5 | Efficacy of surgical face masks in reducing respiratory virus shedding in respiratory droplets and aerosols of symptomatic individuals with seasonal influenza viruses including (a) influenza A and (b) influenza B virus. The figure shows the virus copies per sample collected in nasal swab (red), throat swab (blue), respiratory droplets collected for 30 min while not wearing (dark green) or wearing (light green) a surgical face mask and aerosols collected for 30 min while not wearing (brown) or wearing (orange) a face mask, collected from individuals with acute respiratory symptoms who were positive for influenza A and influenza B virus as determined by RT-PCR in any samples. *P* values for mask intervention as predictor of $\log_{10}$ virus copies per sample in an unadjusted univariate Tobit regression model which allowed for censoring at the lower limit of detection of the RT-PCR assay are shown, with significant differences in bold. For nasal swabs and throat swabs, all infected individuals were included (influenza A virus, n = 31; influenza B virus, n = 14). For respiratory droplets and aerosols, numbers of infected individuals who provided exhaled breath samples while not wearing or wearing a surgical face mask, respectively were: influenza A virus (n = 19 and 19), influenza B virus (n = 6 and 10). A subset of participants provided exhaled breath samples for both mask interventions (influenza A virus, n = 7; influenza B virus, n = 2).



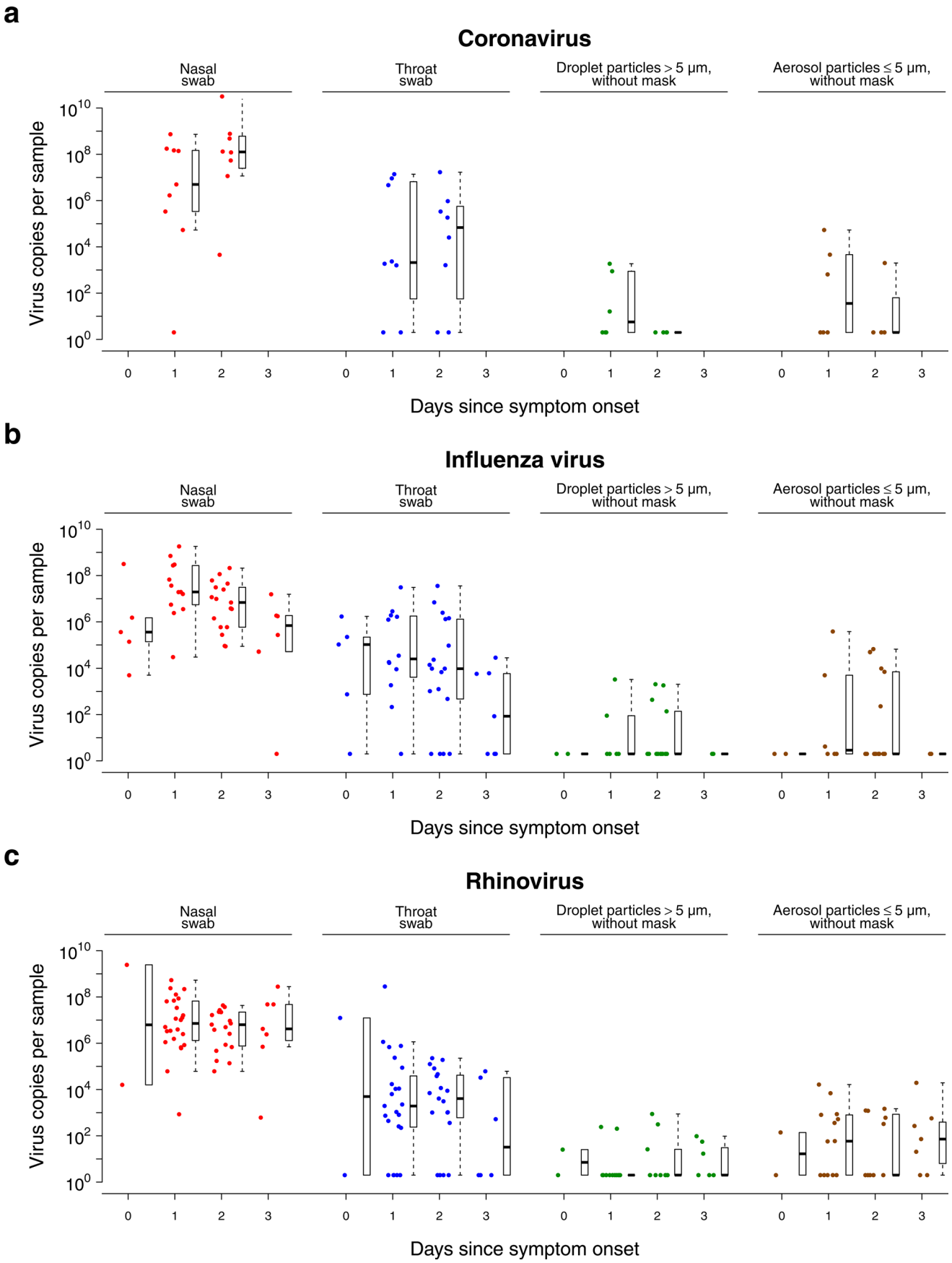
Extended Data Fig. 6 | Correlation of coronavirus viral shedding between different samples (nasal swab, throat swab, respiratory droplets and aerosols) in symptomatic individuals with seasonal coronavirus infection. For nasal swabs and throat swabs, all infected individuals were included (n = 17). For respiratory droplets and aerosols, only infected individuals who provided exhaled breath samples while not wearing a surgical face mask were included (n = 10). r, the Spearman's rank correlation coefficient.



Extended Data Fig. 7 | Correlation of influenza viral shedding between different samples (nasal swab, throat swab, respiratory droplets and aerosols) in symptomatic individuals with seasonal influenza infection. For nasal swabs and throat swabs, all infected individuals were included (n = 43). For respiratory droplets and aerosols, only infected individuals who provided exhaled breath samples while not wearing a surgical face mask were included (n = 23). r, the Spearman's rank correlation coefficient.



Extended Data Fig. 8 | Correlation of rhinovirus viral shedding between different samples (nasal swab, throat swab, respiratory droplets and aerosols) in symptomatic individuals with rhinovirus infection. For nasal swabs and throat swabs, all infected individuals were included (n = 54). For respiratory droplets and aerosols, only infected individuals who provided exhaled breath samples while not wearing a surgical face mask were included (n = 36). r, the Spearman's rank correlation coefficient.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Respiratory virus shedding in respiratory droplets and aerosols stratified by days from symptom onset for (a) coronavirus, (b) influenza virus or (c) rhinovirus. The figures shows the virus copies per sample collected in nasal swab (red), throat swab (blue), respiratory droplets (dark green) and aerosols (brown) collected for 30 min while not wearing a surgical face mask, stratified by the number of days from symptom onset on which the respiratory droplets and aerosols were collected. For nasal swabs and throat swabs, all infected individuals were included (coronavirus, n = 17; influenza virus, n = 43; rhinovirus, n = 54). For respiratory droplets and aerosols, numbers of infected individuals who provided exhaled breath samples while not wearing or wearing a surgical face mask, respectively were: coronavirus (n = 10 and 11), influenza virus (n = 23 and 28), rhinovirus (n = 36 and 32). A subset of participants provided exhaled breath samples for both mask interventions (coronavirus, n = 4; influenza virus, n = 8; rhinovirus, n = 14). The box plots indicate the median with the interquartile range (lower and upper hinge) and ±1.5 × interquartile range from the first and third quartile (lower and upper whisker).

nature research

Corresponding author(s): Benjamin John Cowling

Last updated by author(s): Mar 4, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used.
Data analysis	All analyses were conducted with R version 3.6.0 and the VGAM package 1.1.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Anonymized raw data and R syntax to reproduce all the analyses, figures, tables and supplementary tables in the published article are available at: [Dryad link pending].

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We estimated a priori the sample size to be 300 participants. The primary outcome of the study was the reduction in the exhaled virus concentration of normal tidal breathing by wearing face mask in terms of total virus by RT-PCR as a proxy for infectious virus particle. We expected that a 1-log reduction in exhaled virus particle by face mask intervention would have a clinically relevant effect in reducing the probability of transmission. Except for influenza, there was no quantitative data available from exhaled breath samples from respiratory virus-infected individuals before the present study. If the standard deviation of exhaled virus concentration was 1 log copies/ml (Milton et al., PLoS Pathog 2013), we would detect a difference of >1 log copies/ml in the mask vs control group as long as we have >15 participants with a specific respiratory virus. For example, if our study included 23 participants with rhinovirus detectable in exhaled breath without a mask, we will have 80% power and 0.05 significance level to identify differences in viral shedding in aerosols of 1.28 log10 copies associated with the use of face masks, assuming a standard deviation of 1.54 log10 copies based on data from nasal and throat swab (Lu et al., J Clin Microbiol 2008). We expected from 300 individuals with ARI, at least 150 to have a respiratory virus, and at least 20-30 to have each of rhinovirus, coronavirus, adenovirus and parainfluenza plus small numbers of other respiratory viruses, assuming the Viral Panel would detect respiratory viruses in 60% of participants including 10% by influenza (since we partly recruited during the influenza seasons) and the other 50% made up of rhinovirus, coronavirus, adenovirus and parainfluenza virus.
Data exclusions	As described in the Results section and Supplementary Figure 1, only participants who provided exhaled breath samples and randomized to mask intervention were included; and final analyses were performed only for participants with either coronavirus, influenza virus or rhinovirus infection, which had sufficient sample size for comparison between mask intervention.
Replication	Samples from a subset of participants identified with a coronavirus, influenza or rhinovirus infection were re-tested by RT-PCR with consistent results. R syntax is available to reproduce all the analyses, figures, tables and supplementary tables in the published article.
Randomization	Prior to the exhaled breath collection, each participant was randomly allocated in a 1:1 ratio to either wearing a surgical face mask or not during the exhaled breath collection using a computer-generated sequence. The allocation was concealed to the study stuff performing the exhaled breath collection before allocation of the mask intervention.
Blinding	Blinding to the participant and the study stuff for the mask intervention was not possible. The study staff performing the statistical analyses was also involved in the data collection. We expected there would be minimal bias due to unblinding since data collection for questionnaires was done before randomization to mask intervention, and viral load from a sample measured by RT-PCR is an objective measurement.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Madin-Darby Canine Kidney (MDCK) cells
Authentication	European Collection of Authenticated Cell Cultures.
Mycoplasma contamination	We confirm that all cell lines tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	Nil

Human research participants

Policy information about studies involving human research participants

Population characteristics	As described in the Results section, Table 1a and Supplementary Table 1, there were some differences in characteristics of participants with the different viruses. Overall, most participants were younger adults and 5% were age 11-17 years, but there were more children with influenza virus and no children in the subgroup with coronavirus infection. Overall, 59% were female, but there were more females among the subgroup with coronavirus infection. The majority of participants did not have underlying medical conditions and overall 9% had received influenza vaccination for the current season but only 2% among those with influenza virus infection. The majority of participants were sampled within 24–48 or 48–72 hours of illness onset. 24% of participants had a measured fever ≥37.8ºC, with influenza patients more than twice as likely than coronavirus and rhinovirus-infected patients to have a measured fever. Coronavirus-infected participants coughed the most with an average of 17 (SD 30) coughs during the 30-minute exhaled breath collection. The profile of the participants randomized to with-mask vs without-mask groups were similar.
Recruitment	As described in the Methods section, participants were recruited year-round from March 2013 through May 2016 in a general outpatient clinic of a private hospital in Hong Kong. As routine practice, clinic staff screened all individuals attending the clinics for respiratory and any other symptoms regardless of the purpose of the visit at the triage. Study staff then approached immediately those who reported at least one of the following symptoms of acute respiratory illness (ARI) for further screening: fever≥37.8ºC, cough, sore throat, runny nose, headache, myalgia and phlegm. Individuals who reported ≥2 ARI symptoms, within 3 days of illness onset and ≥11 years of age were eligible to participate.
Ethics oversight	As described in the Methods section, the study protocol was approved by the Institutional Review Board of The University of Hong Kong and the Clinical and Research Ethics Committee of Hong Kong Baptist Hospital.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	The present study was not registered in clinical trials registries, as it was a laboratory-based study of detection of viruses in exhaled breath and the effect of wearing surgical facemasks on virus detection. It was not a Phase II/III clinical trial.
Study protocol	Not available in clinical trials registries (as above). Study protocol will be made available to editors and peer reviewers if requested.
Data collection	As described in the Methods section, participants were recruited year-round from March 2013 through March 2016 in a general outpatient clinic of a private hospital in Hong Kong. Data collection for questionnaires and exhaled breath sample collection was done face-to-face with the participant by trained study staff at the same clinic on the day of participant enrolment.
Outcomes	As pre-specified in the study protocol, the primary outcomes of the study were the virus generation rate in the tidal breathing of participants infected by different respiratory viruses, and the efficacy of face mask in preventing virus dissemination in exhaled breath especially at the aerosol fraction. As pre-specified in the study protocol, one of the secondary outcomes was to provide indirect evidence for relative importance of different transmission routes of influenza and other respiratory viruses. In this regard, in the present manuscript we examined the correlation between viral shedding in nose swabs, throat swabs, respiratory droplets and aerosols, and factors affecting viral shedding in respiratory droplets and aerosols. As described in the Discussion section in the present manuscript about the limitation of our study, there was large proportion of participants with undetectable viral shedding in exhaled breath for each of the viruses studied, and therefore we were unable to examine the exhaled respiratory virus reduction proportion by chi-squared test, nor the exhaled respiratory virus reduction volume (i.e. viral load) by t-test and linear regression as pre-specified in the study protocol. Instead, we have used Fisher's exact test and Tobit regression for the same purposes respectively.

South China Morning Post

Hong Kong / Health & Environment

Coronavirus: hamster research shows effectiveness of masks ‘huge’ in Covid-19 battle, Hong Kong scientists say

- Hamsters placed in adjoining cages with infected subjects were infected at a 66.7 per cent rate; the introduction of a barrier saw the percentage drop to 16.7

drop to 16.7

- ‘It shows very clearly that if infected hamsters or humans … put on masks, they actually protect other people,’ HKU’s Dr Yuen Kwok-yung says

Topic | **Coronavirus pandemic**

Identifying airborne transmission as the dominant route for the spread of COVID-19

Renyi Zhang[a,b,1], Yixin Li[b], Annie L. Zhang[c], Yuan Wang[d], and Mario J. Molina[e,1]

[a]Department of Atmospheric Sciences, Texas A&M University, College Station, TX 77843; [b]Department of Chemistry, Texas A&M University, College Station, TX 77843; [c]Department of Chemistry, College of Natural Sciences, The University of Texas at Austin, Austin, TX 78712; [d]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125; and [e]Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA 92093

Contributed by Mario J. Molina, May 16, 2020 (sent for review May 14, 2020; reviewed by Manish Shrivastava and Tong Zhu)

Various mitigation measures have been implemented to fight the coronavirus disease 2019 (COVID-19) pandemic, including widely adopted social distancing and mandated face covering. However, assessing the effectiveness of those intervention practices hinges on the understanding of virus transmission, which remains uncertain. Here we show that airborne transmission is highly virulent and represents the dominant route to spread the disease. By analyzing the trend and mitigation measures in Wuhan, China, Italy, and New York City, from January 23 to May 9, 2020, we illustrate that the impacts of mitigation measures are discernable from the trends of the pandemic. Our analysis reveals that the difference with and without mandated face covering represents the determinant in shaping the pandemic trends in the three epicenters. This protective measure alone significantly reduced the number of infections, that is, by over 78,000 in Italy from April 6 to May 9 and over 66,000 in New York City from April 17 to May 9. Other mitigation measures, such as social distancing implemented in the United States, are insufficient by themselves in protecting the public. We conclude that wearing of face masks in public corresponds to the most effective means to prevent interhuman transmission, and this inexpensive practice, in conjunction with simultaneous social distancing, quarantine, and contact tracing, represents the most likely fighting opportunity to stop the COVID-19 pandemic. Our work also highlights the fact that sound science is essential in decision-making for the current and future public health pandemics.

COVID-19 | virus | aerosol | public health | pandemic

The novel coronavirus outbreak, coronavirus disease 2019 (COVID-19), which was declared a pandemic by the World Health Organization (WHO) on March 11, 2020, has infected over 4 million people and caused nearly 300,000 fatalities over 188 countries (1). Intensive effort is ongoing worldwide to establish effective treatments and develop a vaccine for the disease. The novel coronavirus, named as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), belongs to the family of the pathogen that is responsible for respiratory illness linked to the 2002–2003 outbreak (SARS-CoV-1) (2). The enveloped virus contains a positive-sense single-stranded RNA genome and a nucleocapsid of helical symmetry of ∼120 nm. There exist several plausible pathways for viruses to be transmitted from person to person. Human atomization of virus-bearing particles occurs from coughing/sneezing and even from normal breathing/talking by an infected person (3–6). These mechanisms of viral shedding produce large droplets and small aerosols (3), which are conventionally delineated at a size of 5 μm to characterize their distinct dispersion efficiencies and residence times in air as well as the deposition patterns along the human respiratory tract (3, 7). Virus transmission occurs via direct (deposited on persons) or indirect (deposited on objects) contact and airborne (droplets and aerosols) routes (3). Large droplets readily settle out of air to cause person/object contamination; in contrast, aerosols are efficiently dispersed in air. While transmission via direct or indirect contact occurs in a short range, airborne transmission via aerosols can occur over an extended distance and time. Inhaled virus-bearing aerosols deposit directly along the human respiratory tract.

Previous experimental and observational studies on interhuman transmission have indicated a significant role of aerosols in the transmission of many respiratory viruses, including influenza virus, SARS-CoV-1, and Middle East Respiratory Syndrome coronavirus (MERS-CoV) (8–11). For example, airborne coronavirus MERS-CoV exhibited strong capability of surviving, with about 64% of microorganisms remaining infectious 60 min after atomization at 25 °C and 79% relative humidity (RH) (9). On the other hand, rapid virus decay occurred, with only 5% survival over a 60-min procedure at 38 °C and 24% RH, indicative of inactivation. Recent experimental studies have examined the stability of SARS-CoV-2, showing that the virus remains infectious in aerosols for hours (12) and on surfaces up to days (12, 13).

Several parameters likely influence the microorganism survival and delivery in air, including temperature, humidity, microbial resistance to external physical and biological stresses, and solar ultraviolet (UV) radiation (7). Transmission and infectivity of airborne viruses are also dependent on the size and number concentration of inhaled aerosols, which regulate the amount (dose) and pattern for respiratory deposition. With typical nasal breathing (i.e., at a velocity of ∼1 $m{\cdot}s^{-1}$) (4), inhalation of airborne viruses leads to direct and continuous deposition into the human respiratory tract. In particular, fine aerosols (i.e., particulate

Significance

We have elucidated the transmission pathways of coronavirus disease 2019 (COVID-19) by analyzing the trend and mitigation measures in the three epicenters. Our results show that the airborne transmission route is highly virulent and dominant for the spread of COVID-19. The mitigation measures are discernable from the trends of the pandemic. Our analysis reveals that the difference with and without mandated face covering represents the determinant in shaping the trends of the pandemic. This protective measure significantly reduces the number of infections. Other mitigation measures, such as social distancing implemented in the United States, are insufficient by themselves in protecting the public. Our work also highlights the necessity that sound science is essential in decision-making for the current and future public health pandemics.

Author contributions: R.Z. designed research; R.Z., Y.L., and Y.W. performed research; R.Z., Y.L., Y.W., and M.J.M. analyzed data; and R.Z., A.L.Z., and M.J.M. wrote the paper.

Reviewers: M.S., Pacific Northwest National Laboratory; and T.Z., Peking University.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: renyi-zhang@tamu.edu or mjmolina@ucsd.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2009637117/-/DCSupplemental.

First published June 11, 2020.

matter smaller than 2.5 μm, or $PM_{2.5}$) penetrate deeply into the respiratory tract and even reach other vital organs (14, 15). In addition, viral shedding is dependent on the stages of infection and varies between symptomatic and asymptomatic carriers. A recent finding (16) showed that the highest viral load in the upper respiratory tract occurs at the symptom onset, suggesting the peak of infectiousness on or before the symptom onset and substantial asymptomatic transmission for SARS-CoV-2.

The COVID-19 outbreak is significantly more pronounced than that of the 2002/2003 SARS, and the disease continues to spread at an alarming rate worldwide, despite extreme measures taken by many countries to constrain the pandemic (1). The enormous scope and magnitude of the COVID-19 outbreak reflect not only a highly contagious nature but also exceedingly efficient transmission for SARS-CoV-2. Currently, the mechanisms to spread the virus remain uncertain (17), particularly considering the relative contribution of the contact vs. airborne transmission routes to this global pandemic. Available epidemiological (1) and experimental (12, 18) evidence, however, implicates airborne transmission of SARS-CoV-2 via aerosols as a potential route for the spreading of the disease.

Distinct Pandemic Trends in the Three Epicenters

To gain insight into the mechanism of the virus transmission routes and assess the effectiveness of mitigation measures, we analyzed the trend of the pandemic worldwide from January 23 to May 9, 2020 (Fig. 1). The COVID-19 outbreak initially emerged during December 2019 in Wuhan, China (1). The numbers of confirmed infections and fatalities in China dominated the global trend during January and February 2020 (Fig. 1*A*), but the increases in the newly confirmed cases and fatalities in China have exhibited sharp declines since February (Fig. 1*B*). In contrast to the curve flattening in China, those numbers in other countries have increased sharply since the beginning of March. The epicenter shifted from Wuhan to Italy in early March and to New York City (NYC) in early April. By April 30, the numbers of confirmed COVID-19 cases and deaths, respectively, reached over 200,000 and 27,000 in Italy and over 1,000,000 and 52,000 in the United States, compared to about 84,000 and 4,600 in China (Fig. 1*B*). Notably, the curves in Italy exhibit a slowing trend since mid-April, while the numbers in the world and the United States continue to increase. Remarkably, the recent trends in the numbers of infections and fatalities in the world and in the United States exhibit striking linearity since the beginning of April (Fig. 1*C*).

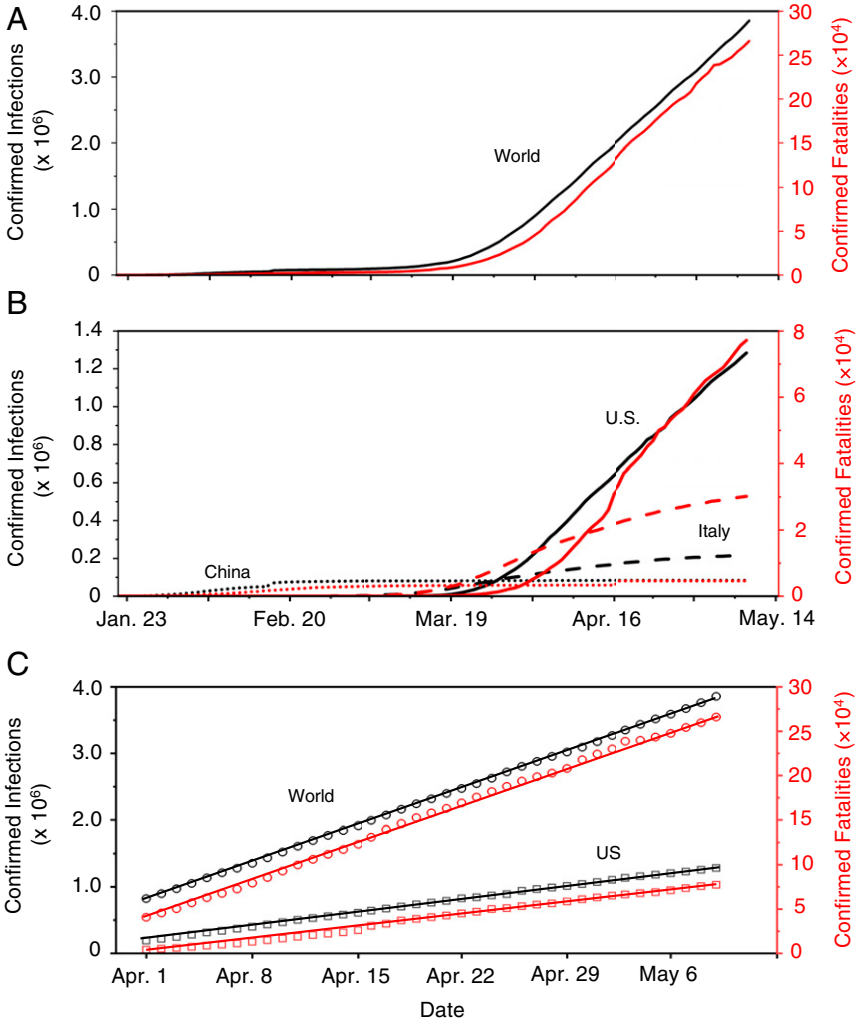


Fig. 1. Distinct global trends of the COVID-19 pandemic. (*A*) Confirmed infections and fatalities worldwide. (*B*) Comparison of the confirmed infections and fatalities between China, Italy, and United States. (*C*) Linear regression of the confirmed infections and fatalities worldwide and in United States from April 1 to May 9, 2020; the linear regression is, respectively, $y = 79{,}398x + 810{,}167$ ($R^2 = 0.999$) for infections and $y = 6{,}075x + 39{,}409$ ($R^2 = 0.998$) for fatalities worldwide and $y = 28{,}971x + 201{,}187$ ($R^2 = 0.999$) for infections and $y = 2{,}059x + 243$ ($R^2 = 0.995$) for fatalities in the United States. The left axis and black color correspond to the numbers of confirmed infections, and the right axis and red color represent the confirmed fatalities.

We interpreted the differences in the pandemic trends by considering the mitigation measures implemented worldwide. The curve flattening in China can be attributed to extensive testing, quarantine, and contact tracing; other aggressive measures implemented in China include lockdown of all cities and rural areas in the whole country, isolation of residents having close contact with infected people, and mandated wearing of face masks in public. However, the effectiveness of those mitigation measures has yet to be rigorously evaluated. Differentiation of the effects of those mitigation measures in China is challenging (19), since the implementation occurred almost simultaneously in January 2020. While similar quarantine, isolation, and city lockdown measures were also implemented on March 9 in Italy after the country became the second epicenter, the curve of infections has yet to show complete flattening. In the United States, guidelines for social distancing, quarantine, and isolation were issued by the federal government on March 16, and stay-at-home orders were implemented by many state and local governments starting, for example, on March 19 and April 3 and on March 22 in NYC. The social distancing measures implemented in the United States include staying at least 6 feet (~2 m) away from other people, no gathering in groups, staying out of crowded places, and avoiding mass gatherings (20). Obviously, the continuous rise in the US infected numbers casts doubt on the effectiveness of those preventive measures alone (Fig. 1 *B* and *C*).

In contrast to China, wearing of face masks was not mandated and was unpopular in most of the western world during the early outbreak of the pandemic. Advice on the use of face masks was not issued until April 6, 2020 by the WHO (1), claiming that it is important only to prevent infected persons from viral transmission by filtering out droplets but that it is unimportant to prevent uninfected persons from breathing virus-bearing aerosols. The regions heavily plagued by COVID-19 in northern Italy, such as Lombard, ordered face covering in public starting on April 6, and the Italian authorities required nationwide mandatory use of face masks on May 4. All New Yorkers were mandated to use face covering in public starting on April 17, when social distancing was not possible. With measures implemented in the United States seemingly comparable to those in China, social distancing, quarantine, and isolation exhibited little impact on stopping the spreading of the disease in the United States, as reflected by the linearity from April 1 to May 9 (Fig. 1*C*). It is possible, however, that these measures likely alter the slope of the infection curve, that is, by reducing the rate of infections during the early stage of the pandemic (Fig. 1). Notably, the recommended physical separation for social distancing is beneficial to prevent direct contact transmission but is insufficient (without face masks) to protect inhalation of virus-bearing aerosols (or even small droplets at intermediate proximity), owing to rapid air mixing (7).

Understanding the Impacts of Face Covering

Compared to the simultaneous implementation of measures in China, intervention measures were successively implemented in the western world (Fig. 2*A*), providing an opportunity for assessing their relative effectiveness. We quantified the effects of face covering by projecting the number of infections based on the data prior to implementing the use of face masks in Italy on April 6 and NYC on April 17 (Fig. 2*A*; see *Methods*). Such projections are reasonable considering the excellent linear correlation for the data prior to the onset of mandated face covering (Fig. 2 *B* and *C* and *SI Appendix*, Fig. S1). Our analysis indicates that face covering reduced the number of infections by over 78,000 in Italy from April 6 to May 9 and by over 66,000 in NYC from April 17 to May 9. In addition, varying the correlation from 15 d to 30 d prior to the onset of the implementation reveals little difference in the projection for both places, because of the high correlation coefficients (*SI Appendix*, Fig. S1). Notably, the trends of the infection curves in Italy and NYC contrast to those in the world and in the United States (Fig. 1*C*), which show little deviation from the linearity due to the nonimplementation of face-covering measures globally and nationally, respectively. The inability of social distancing, quarantine, and isolation alone to curb the spread of COVID-19 is also evident from the linearity of the infection curve prior to the onset of the face-covering rule in Italy on April 6 and in NYC on April 17 (Fig. 2 *B* and *C*). Hence, the difference made by implementing face covering significantly shapes the pandemic trends worldwide.

We further compared the numbers of daily new cases between NYC and the United States (excluding the data in NYC) from March 1 to May 9 (Fig. 3). The daily numbers of newly confirmed infections in NYC and the United States show a sharp increase in late March and early April. There exists a slower increase in the number after implementation of the stay-at-home order (about 14 d in New York and shortly after April 3 in the United States), which is attributable to the impacts of this measure. After April 3, the only difference in the regulatory measures between NYC and the United States lies in face covering on April 17 in NYC. We applied linear regression to the data between April 17 and May 9 in NYC and between April 5 and May 9 in the United States. While the daily numbers of newly confirmed infections fluctuate considerably, the slope of the regression unambiguously reflects the trend in both data. The daily new infection in NYC decreases with a slope of 106 cases per day after April 17, corresponding to a decreasing rate of ~3% per day (relative to April 17). For comparison, the daily new infections in the United States (excluding NYC) increase, with a slope of 70 cases per day after April 4, corresponding to an increasing rate of ~0.3% per day (relative to April 5). Hence, the decreasing rate in the daily new infections in NYC with mandated face covering is in sharp contrast to that in the United States with only social-distancing and stay-at-home measures, further confirming the importance of face covering in intervening the virus transmission.

Dominant Airborne Transmission

We further elucidated the contribution of airborne transmission to the COVID-19 outbreak by comparing the trends and mitigation measures during the pandemic worldwide and by considering the virus transmission routes (Fig. 4). Face covering prevents both airborne transmission by blocking atomization and inhalation of virus-bearing aerosols and contact transmission by blocking viral shedding of droplets. On the other hand, social distancing, quarantine, and isolation, in conjunction with hand sanitizing, minimize contact (direct and indirect) transmission but do not protect against airborne transmission. With social distancing, quarantine, and isolation in place worldwide and in the United States since the beginning of April, airborne transmission represents the only viable route for spreading the disease, when mandated face covering is not implemented. Similarly, airborne transmission also contributes dominantly to the linear increase in the infection prior to the onset of mandated face covering in Italy and NYC (Fig. 2 *B* and *C* and *SI Appendix*, Fig. S1). Hence, the unique function of face covering to block atomization and inhalation of virus-bearing aerosols accounts for the significantly reduced infections in China, Italy, and NYC (Figs. 1–3), indicating that airborne transmission of COVID-19 represents the dominant route for infection.

Recent measurements identified SARS-Cov-2 RNA on aerosols in Wuhan's hospitals (18) and outdoor in northern Italy (21), unraveling the likelihood of indoor and outdoor airborne transmission. Within an enclosed environment, virus-bearing aerosols from human atomization are readily accumulated, and elevated levels of airborne viruses facilitate transmission from person to person. Transmission of airborne viruses in open air is subject to

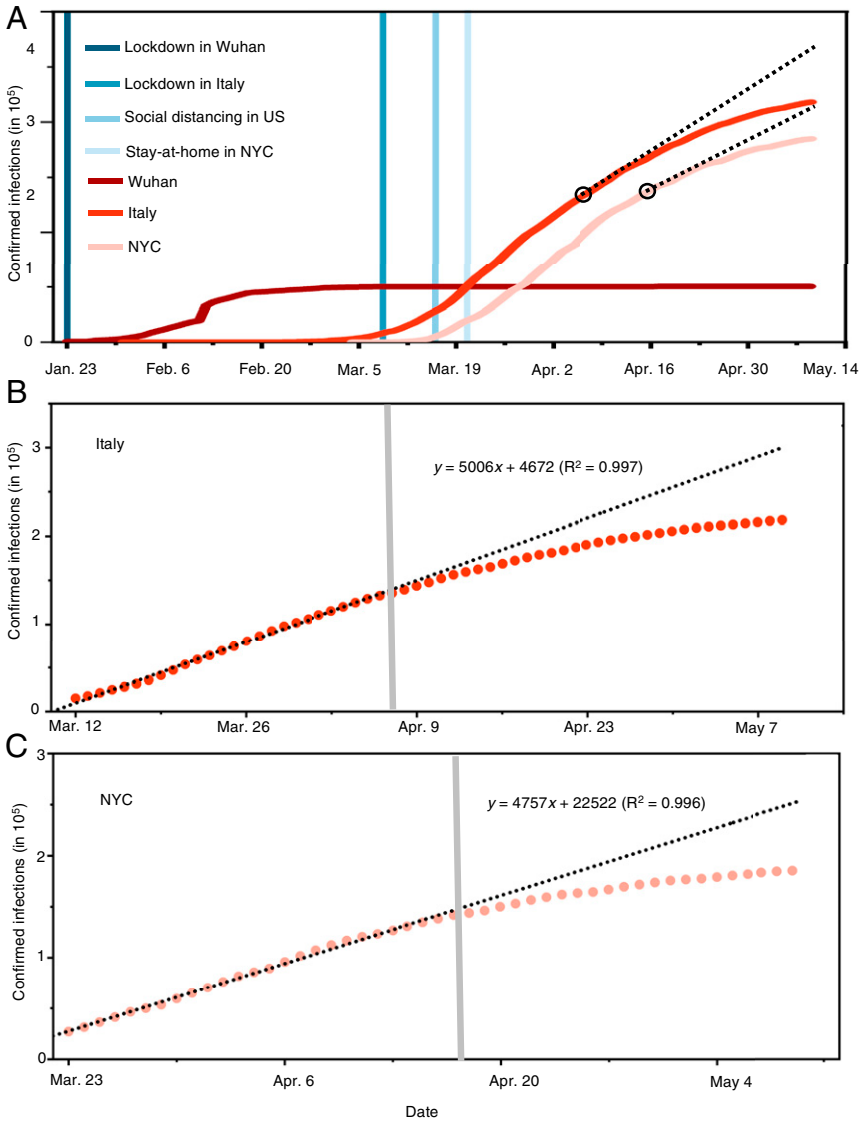


Fig. 2. The evolving epicenter from Wuhan, to Italy, to NYC. (*A*) Comparison of the trends and mitigation measures between Wuhan, Italy, and NYC in 2020. The vertical lines mark the date for implementing mitigation measures. The two black circles label the dates when face covering was implemented: April 6 in northern Italy and April 17 in NYC. The black dashed lines represent the projection without face covering based on linear regression of 26-d data prior to implementing this measure. (*B*) Linear regression of the number of confirmed infections for 26-d data prior to implementing face covering in Italy. The shaded vertical line denotes the date when face covering was implemented on April 6 in northern Italy. (*C*) Linear regression of the number of confirmed infections for 26-d data prior to implementing face covering in NYC. The shaded vertical line denotes the date when face covering was implemented on April 17 in NYC. In *B* and *C*, the circles are reported values, and the dotted line represents fitting and projection of the confirmed infections before and after face-covering, respectively.

dilution, although virus accumulation still occurs due to stagnation under polluted urban conditions (7, 22). Removal of virus-bearing particles from human atomization via deposition is strongly size dependent, with the settling velocities ranging from 2.8×10^{-5} m·s^{-1} to 1.4×10^{-3} m·s^{-1} for the sizes of 1 and 10 μm, respectively (7). For comparison, typical wind velocity is about 1 m·s^{-1} to 3 m·s^{-1} indoors (23) and is ∼1 m·s^{-1} horizontally and 0.1 m·s^{-1} vertically in stable air (7, 22). Under those indoor and outdoor conditions, the residence time of virus-bearing aerosols reaches hours, due to air mixing (7).

We also examined ambient conditions relevant to the outbreaks in Wuhan, Italy, and NYC. The initial outbreak of COVID-19 in Wuhan coincided with the winter haze season in China (7, 22), during which high levels of $PM_{2.5}$ were prevalent in air (*SI Appendix*, Figs. S2 and S3). On the other hand, the daily average $PM_{2.5}$ concentrations were much lower during the outbreaks in Rome, Italy, and in NYC (*SI Appendix*, Fig. S2). The airborne transmission pathways (i.e., indoor or outdoor) as well as the effects of ambient $PM_{2.5}$ levels on virus transmission may be variable among urban cities. For example, the winter haze conditions in China likely exacerbated outdoor virus spreading (24, 25), because of low UV radiation, air stagnation (lacking ventilation on the city scale), and low temperature (7, 22). Also, there may exist a synergetic effect of simultaneous exposure to the virus and $PM_{2.5}$ to enhance the infectivity, severity, and fatalities of the disease (14, 26). In addition, nascent virus-bearing aerosols produced from human atomization likely undergo transformation in air, including coagulation with ambient preexisting PM and/or growth on a time scale of a few hours in typical urban air (27–29). Such transformation, as recently documented on coarse PM in Italy (21), may mitigate

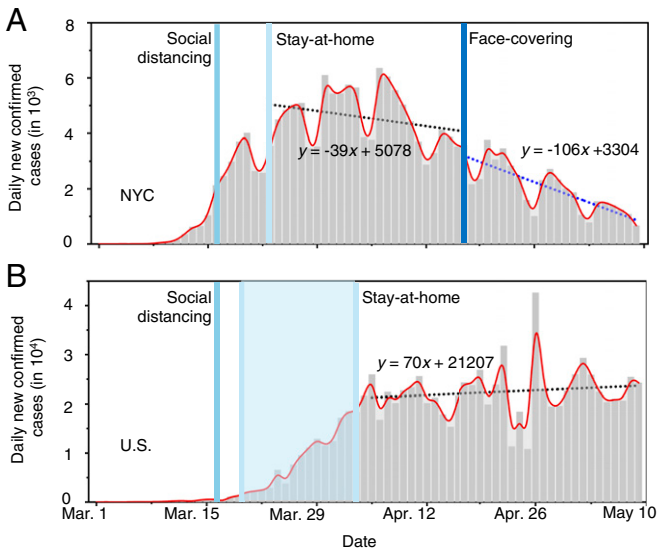


Fig. 3. Contrasting the trends of new infections between NYC and the United States. Daily new confirmed infections in (*A*) NYC and (*B*) the United States. The dotted lines represent linear fitting to the data between April 17 and May 9 in NYC and between April 4 and May 9 in the United States. In *B*, the number in NYC was subtracted from that in the United States. The vertical lines label the dates for social distancing, stay-at-home orders, and mandated face-covering.

virus inactivation (9, 12), by providing a medium to preserve its biological properties and elongating its lifetimes. However, key questions remain concerning transformation and transmission of virus-bearing aerosols from human atomization in air. Specifically, what are the impacts of transformation of human-atomized aerosols on viral surviving and infectivity in air?

While the humidity effect on viral surviving is uncertain (3, 9), the conditions during the outbreaks in Wuhan, Rome, and NYC correspond to high RH yet low absolute humidity because of low temperature (*SI Appendix*, Fig. S3). Early experimental work (9) showed remarkable survival for the analogous coronavirus MERS-CoV at the RH level characteristic of the COVID-19 outbreaks in Wuhan, Rome, and NYC. For comparison, indoor temperature and RH typically range from 21 °C to 27 °C and 20 to 70%, respectively (23).

Of particular importance are the considerations that render airborne SARS-CoV-2 the most efficient among all transmission routes. Even with normal nasal breathing, inhalation of virus-bearing aerosols results in deep and continuous deposition into the human respiratory tract, and this transmission route typically requires a low dose (8). Also, airborne viruses have great mobility and sufficiently long surviving time for dispersion (9, 12), and residents situated in densely populated environments are highly vulnerable. In addition, nascent micrometer-size aerosols produced from coughing/sneezing of infected people have the potential of containing many viruses, particularly for asymptomatic carriers (16).

Future research is critically needed to assess the transmission, transformation, and dispersion of virus-bearing aerosols from human atomization under different environmental conditions, as well as the related impacts on virus infectivity. It is equally important to understand human atomization of airborne viruses: What are the number and size distributions of nascent aerosols as well as the viral load per particle from coughing/sneezing? It is also imperative to evaluate human inhalation of airborne viruses: How are aerosols deposited along the respiratory tract, and what is the minimum dose of airborne viruses required for infection? It is also important to evaluate the performance of face masks to quantify the efficiency to filtrate airborne viruses relevant to human atomization and inhalation. Elucidation of these mechanisms requires an interdisciplinary effort.

A Policy Perspective

The governments' responses to the COVID pandemic have so far differed significantly worldwide. Swift actions to the initial outbreak were undertaken in China, as reflected by nearly simultaneous implementation of various aggressive mitigation measures. On the other hand, the response to the pandemic was generally slow in the western world, and implementation of the intervention measures occurred only consecutively. Clearly, the responsiveness of the mitigation measures governed the evolution, scope, and magnitude of the pandemic globally (Figs. 1 and 2).

Curbing the COVID-19 relies not only on decisive and sweeping actions but also, critically, on the scientific understanding of the virus transmission routes, which determines the effectiveness of the mitigation measures (Fig. 5). In the United States, social distancing and stay-at-home measures, in conjunction with hand sanitizing (Fig. 5, path a), were implemented during the early stage of the pandemic (March 16) (20). These measures minimized short-range contact transmission but did not prevent long-range airborne transmission, responsible for the inefficient containing of the pandemic in the United States (Figs. 1 and 3). Mandated face covering, such as those implemented in China, Italy, and NYC, effectively prevented airborne transmission by blocking atomization and inhalation of virus-bearing aerosols and contact transmission by blocking viral shedding of droplets. While the combined face-covering and social distancing measures offered dual protection

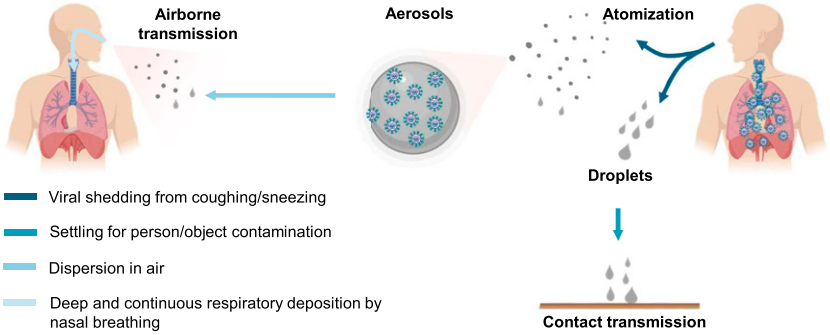


Fig. 4. Transmission of COVID-19. Human atomization of viruses arises from coughing or sneezing of an infected person, producing virus-containing droplets (>5 μm) and aerosols (<5 μm). Virus transmission from person to person occurs through direct/indirect contact and airborne aerosol/droplet routes. Large droplets mainly settle out of air to cause person/object contamination, while aerosols are efficiently dispersed in air. Direct and airborne transmissions occur in short range and extended distance/time, respectively. Inhaled airborne viruses deposit directly into the human respiration tract.

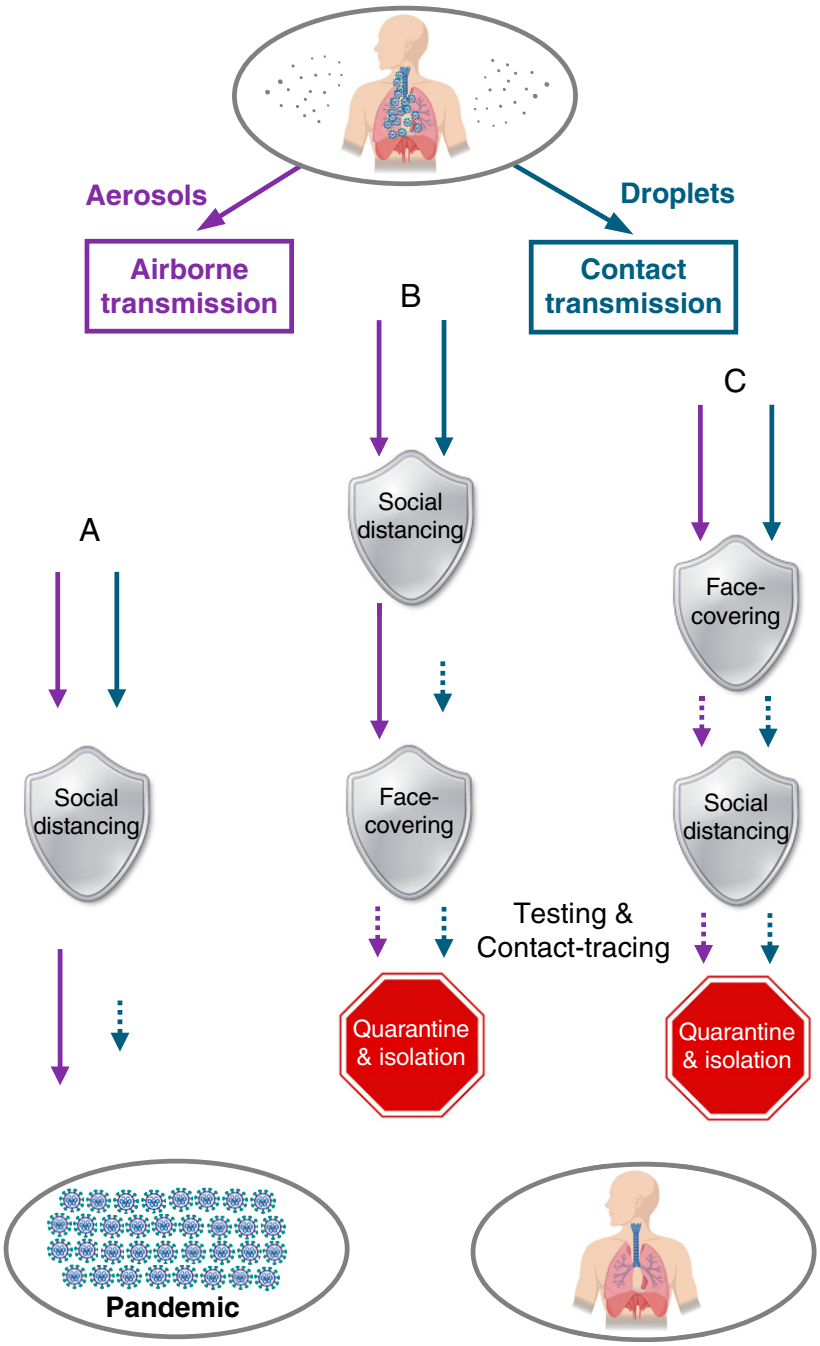


Fig. 5. Mitigation paradigm. Scenarios of virus transmission under the distancing/quarantine/isolation measure only (path a), the measures with distancing/quarantine/isolation followed by face covering (path b), and the measures with simultaneous face covering and distancing/quarantine/isolation (path c). The short-dashed arrows label possible remnants of virus transmission due to circumstances when the measure is not possible or disobeyed and/or imperfection of the measure.

against the virus transmission routes, the timing and sequence in implementing the measures also exhibited distinct outcomes during the pandemic. For example, social distancing measures, including city lockdown and stay-at-home orders, were implemented well before face covering was mandated in Italy and NYC (Fig. 5, path b), and this sequence left an extended window (28 d in Italy and 32 d in NYC) for largely uninterrupted airborne transmission to spread the disease (Figs. 2 and 3). The simultaneous implementation of face covering and social distancing (Fig. 5, path c), such as that undertaken in China, was most optimal, and this configuration, in conjunction with extensive testing and contact tracing, was responsible for the curve flattening in China (Fig. 1). Also, there likely existed remnants of virus transmission after the implementation of regulatory measures, because of circumstances when the measures were not practical or were disobeyed and/or imperfection of the measures. Such limitations, which have been emphasized by the WHO (1), spurred on controversial views on the validity of wearing face masks to prevent the virus transmission during the pandemic (30). However, it is implausible that the limitations of mitigation measures alone contributed dominantly to the global pandemic trend, as exemplified by the success in China. Our work suggests that the failure in containing the propagation of COVID-19 pandemic worldwide is largely attributed to the unrecognized importance of airborne virus transmission (1, 20).

Conclusions

The inadequate knowledge on virus transmission has inevitably hindered development of effective mitigation policies and resulted in unstoppable propagation of the COVID-19 pandemic (Figs. 1–3). In this work, we show that airborne transmission, particularly via nascent aerosols from human atomization, is highly virulent and represents the dominant route for the transmission of this disease. However, the importance of airborne transmission has not been considered in establishment of mitigation measures by government authorities (1, 20). Specifically, while the WHO and the US Centers for Disease Control and Prevention (CDC) have emphasized the prevention of contact transmission, both WHO and CDC have largely ignored the importance of the airborne transmission route (1, 20). The current mitigation measures, such as social distancing, quarantine, and isolation implemented in the United States, are insufficient by themselves in protecting the public. Our analysis reveals that the difference with and without mandated face covering represents the determinant in shaping the trends of the pandemic worldwide. We conclude that wearing of face masks in public corresponds to the most effective means to prevent interhuman transmission, and this inexpensive practice, in conjunction with extensive testing, quarantine, and contact tracking, poses the most probable fighting opportunity to stop the COVID-19 pandemic, prior to the development of a vaccine. It is also important to emphasize that sound science should be effectively communicated to policy makers and should constitute the prime foundation in decision-making amid this pandemic. Implementing policies without a scientific basis could lead to catastrophic consequences, particularly in light of attempts to reopen the economy in many countries. Clearly, integration between science and policy is crucial to formulation of effective emergency responses by policy makers and preparedness by the public for the current and future public health pandemics.

Methods

Projection of the pandemic trend without implementing face covering in Italy and NYC was performed first by establishing the linear correlation between the infection number and date. We considered the data for both 15 and 30 d prior to the onset of face covering (*SI Appendix*, Fig. S1). The slope and the reported infection number were used for the projections. The avoided infection number due the face covering was determined from the difference between the projected and reported values on May 9, 2020.

The data for accumulative confirmed infections and fatalities in Wuhan, Italy, and NYC were taken from the reports by Wuhan Municipal Health Commission (http://wjw.wuhan.gov.cn/), European CDC (https://www.ecdc.europa.eu/en), and NYC government (https://www1.nyc.gov/site/doh/covid/covid-19-data.page), respectively. The data of accumulative confirmed infections and fatalities worldwide were taken from WHO COVID-19 situation report (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports) (1), and the numbers in China, Italy, and United States were from taken from European CDC.

Ground-based measurements of $PM_{2.5}$ and RH in Wuhan were taken from the China National Environmental Monitoring Centre (http://beijingair.sinaapp.com/). The $PM_{2.5}$ data in NYC were taken from US Environmental Protection Agency (https://www.epa.gov/outdoor-air-quality-data). The $PM_{2.5}$ data in Rome were taken were from Centro Regionale della Qualità dell'aria (http://www.arpalazio.net/main/aria/). The RH data in Rome and NYC were taken from the 6-hourly interim reanalysis of the European Centre for Medium-range Weather Forecasts (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5).

We used spaceborne measurements of aerosol optical depth (AOD) to characterize the regional aerosol pollution during the COVID-19 outbreak (January 23 to February 10, 2020) in China. The green band AODs at 0.55 μm are available from Terra and Aqua combined Moderate Resolution Imaging Spectroradiometer Version 6 Multiangle Implementation of Atmospheric Correction (https://lpdaac.usgs.gov/products/mcd19a2v006/). The Level-2 product has daily global coverage with 1-km pixel resolution. The AOD retrieval is only available for the clear sky.

Data Availability. All data relevant to this research are available in the main text and *SI Appendix*.

ACKNOWLEDGMENTS. This work was supported by the Robert A. Welch Foundation (Grant A-1417). A.L.Z. acknowledges the support of a fellowship from the Robert A. Welch Foundation. We are grateful to Fang Zhang for the $PM_{2.5}$ data in Wuhan, China.

1. World Health Organization, Coronavirus disease (COVID-2019) situation reports. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/. Accessed 9 May 2020.
2. A. R. Fehr, S. Perlman, Coronaviruses: An overview of their replication and pathogenesis. *Methods Mol. Biol.* **1282**, 1–23 (2015).
3. J. S. Kutter, M. I. Spronken, P. L. Fraaij, R. A. Fouchier, S. Herfst, Transmission routes of respiratory viruses among humans. *Curr. Opin. Virol.* **28**, 142–151 (2018).
4. J. W. Tang *et al*., Airflow dynamics of human jets: Sneezing and breathing - potential sources of infectious aerosols. *PLoS One* **8**, e59970 (2013).
5. N. H. L. Leung *et al*., Respiratory virus shedding in exhaled breath and efficacy of face masks. *Nat. Med.* **26**, 676–680 (2020).
6. V. Stadnytskyi, C. E. Bax, A. Bax, P. Anfinrud, The airborne lifetime of small speech droplets and their potential importance in SARS-CoV-2 transmission. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 11875–11877 (2020).
7. R. Zhang *et al*., Formation of urban fine particulate matter. *Chem. Rev.* **115**, 3803–3855 (2015).
8. R. Tellier, Aerosol transmission of influenza A virus: A review of new studies. *J. R. Soc. Interface* **6** (suppl. 6), S783–S790 (2009).
9. O. V. Pyankov, S. A. Bodnev, O. G. Pyankova, I. E. Agranovski, Survival of aerosolized coronavirus in the ambient air. *J. Aerosol Sci.* **115**, 158–163 (2018).
10. M. Richard, R. A. M. Fouchier, Influenza A virus transmission via respiratory aerosols or droplets as it relates to pandemic potential. *FEMS Microbiol. Rev.* **40**, 68–85 (2016).
11. T. P. Weber, N. I. Stilianakis, Inactivation of influenza A viruses in the environment and modes of transmission: A critical review. *J. Infect.* **57**, 361–373 (2008).
12. N. van Doremalen *et al*., Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N. Engl. J. Med.* **382**, 1564–1567 (2020).
13. A. W. H. Chin *et al*., Stability of SARS-CoV-2 in different environmental conditions. *Lancet* **1**, E10 (2020).
14. K. A. Rychlik *et al*., In utero ultrafine particulate matter exposure causes offspring pulmonary immunosuppression. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 3443–3448 (2019).
15. G. Wu *et al*., Adverse organogenesis and predisposed long-term metabolic syndrome from prenatal exposure to fine particulate matter. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11590–11595 (2019).
16. X. He *et al*., Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 672–675 (2020).
17. D. Lewis, Is the coronavirus airborne? Experts can't agree. *Nature* **580**, 175 (2020).
18. Y. Liu *et al*., Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. *Nature*, 10.1038/s41586-020-2271-3 (2020).
19. L. Ferretti *et al*., Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, eabb6936 (2020).
20. US Centers for Disease Control and Prevention, Coronavirus Disease 2019 (COVID-19) - Social distancing, quarantine, and isolation. https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html. Accessed 9 May 2020.
21. L. Setti *et al*, SARS-Cov-2 RNA found on particulate matter of Bergamo in Northern Italy: First preliminary evidence. *Environ. Res.*, 10.1016/j.envres.2020.109754 (2020).0013-9351
22. Z. An *et al*., Severe haze in northern China: A synergy of anthropogenic emissions and atmospheric processes. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8657–8666 (2019).
23. L. A. Wallace, S. J. Emmerich, C. Howard-Reed, Continuous measurements of air change rates in an occupied house for 1 year: The effect of temperature, wind, fans, and windows. *J. Expo. Anal. Environ. Epidemiol.* **12**, 296–306 (2002).
24. Q. Ye, J. F. Fu, J. H. Mao, S. Q. Shang, Haze is a risk factor contributing to the rapid spread of respiratory syncytial virus in children. *Environ. Sci. Pollut. Res. Int.* **23**, 20178–20185 (2016).
25. Z. Gong *et al*., Probable aerosol transmission of severe fever with thrombocytopenia syndrome virus in southeastern China. *Clin. Microbiol. Infect.* **21**, 1115–1120 (2015).
26. X. Wu *et al*., Exposure to air pollution and COVID-19 mortality in the United States. https://projects.iq.harvard.edu/files/covid-pm/files/pm_and_covid_mortality.pdf. Accessed 9 May 2020.
27. S. Guo *et al*., Elucidating severe urban haze formation in China. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17373–17378 (2014).
28. F. Zhang *et al*., An unexpected catalyst dominates formation and radiative forcing of regional haze. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 3960–3966 (2020).
29. J. Peng *et al*., Markedly enhanced absorption and direct radiative forcing of black carbon under polluted urban environments. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4266–4271 (2016).
30. J. Howard *et al*, Face masks against COVID-19: An evidence review. https://doi.org/10.20944/preprints202004.0203.v2 (13 May 2020).

Annals of Internal Medicine

Sexual Health in the SARS-CoV-2 Era

Jack L. Turban, MD, MHS; Alex S. Keuroghlian, MD, MPH; and Kenneth H. Mayer, MD

More than 200 000 people have died of severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) infection, leading to widespread concern regarding physical morbidity and mortality. The sexual health implications, however, have received little focus. On the basis of existing data, it appears all forms of in-person sexual contact carry risk for viral transmission, because the virus is readily transmitted by aerosols and fomites. This has resulted in broad guidance regarding physical distancing, with substantial implications for sexual well-being. Given the important role of sexuality in most people's lives, health care providers (HCPs) should consider counseling patients on this topic whenever possible. This is an unprecedented and stressful time for HCPs; facilitating brief conversations and referrals to relevant resources (**Table**) can help patients maintain sexual wellness amid the pandemic.

Current Evidence Suggests That All In-person Sexual Contact Carries Transmission Risk

SARS-CoV-2 is present in respiratory secretions and spreads through aerosolized particles (1). It may remain stable on surfaces for days (1). On the basis of this information, all types of in-person sexual activity probably carry risk for SARS-CoV-2 transmission. Infected individuals have the potential to spread respiratory secretions onto their skin and personal objects, from which the virus can be transmitted to a sexual partner. Because many SARS-CoV-2–infected people are asymptomatic, HCPs are left with little to offer beyond guidance to not engage in any in-person sexual activity.

Data are lacking regarding other routes of sexual transmission. Two small studies of SARS-CoV-2–infected people did not detect virus in semen or vaginal secretions (2, 3). An additional study of semen samples from 38 patients detected the virus by reverse transcriptase polymerase chain reaction in 6 patients (15.8%) (4). However, the relevance regarding sexual transmission remains unknown. Until this is better understood, it would be prudent to consider semen potentially infectious. Although 1 study failed to detect the virus in urine samples (5), there is evidence that SARS-CoV-2 nucleic acids were detected in a urine sample in at least 1 patient in another study (6). Until this is clarified, urine should also be considered potentially infectious. SARS-CoV-2 RNA has been detected in stool samples, raising concern for fecal–oral transmission (7). It is not clear, however, whether viral RNA detected in stool is capable of causing productive infection. Moreover, these data are moot, given that any in-person contact results in substantial risk for disease transmission owing to the virus' stability on common surfaces and propensity to propagate in the oropharynx and respiratory tract.

Psychological Effects of Sexual Abstinence

Sexual expression is a central aspect of human health but is often neglected by HCPs. Messaging around sex being dangerous may have insidious psychological effects at a time when people are especially susceptible to mental health difficulties. Some groups, including sexual and gender minority (SGM) communities, may be particularly vulnerable to sexual stigma, given the historical trauma of other pandemics, such as AIDS. Abstinence recommendations may conjure memories of the widespread stigmatization of SGM people during the AIDS crisis. For the population at large, a recommendation of long-term sexual abstinence is unlikely to be effective, given the well-documented failures of abstinence-based public health interventions and their likelihood to promote shame (8).

HCPs Should Consider Counseling on Safe Sexual Practices and Risk Reduction Whenever Possible

A range of sexual practices organized from least to most risky is shown in the **Table**. Abstinence is the lowest-risk approach to sexual health during the pandemic. Masturbation is an additional safe recommendation for patients to meet their sexual needs without the risk for SARS-CoV-2 infection.

Given that abstinence-only recommendations, however, are likely to promote shame and unlikely to achieve intended behavioral outcomes (8), sex-positive recommendations regarding remote sexual activity are optimal during the pandemic, balancing human needs for intimacy with personal safety and pandemic control. Patients can be counseled to engage in sexual activity with partners via the telephone or video chat services. Given privacy concerns, they should be counseled to use secure encrypted platforms. They should also be warned about the risks for sexual partners taking screenshots of conversations and relevant risks and laws regarding sexual extortion. For some patients, including those without internet access and minors at home from school who are in environments unaccepting of their sexual orientation, digital sexual practices may not be feasible. During all conversations, HCPs should express a nonjudgmental stance to encourage comfortable discussion and minimize shame. This is particularly important with minors, because fear of

This article was published at Annals.org on 8 May 2020.

Table. Sexual Practices During the SARS-CoV-2 Era and Patient Resources

Sexual Approach	Summary
Sexual abstinence	Low risk for infection, though not feasible for many
Masturbation	Low risk for infection Safe masturbation tips (Planned Parenthood): https://www.plannedparenthood.org/learn/sex-pleasure-and-sexual-dysfunction/masturbation
Sexual activity via digital platforms, such as the phone or video chat	Patients should be counseled on the risk for screenshots of conversations or videos and sexual extortion Minors should be counseled on potential legal consequences if they are in possession of sexual images of other minors Minors should be counseled on the risks for online sexual predation, which has increased since the pandemic began Speaking with children about sexual risk online during COVID-19 (Scientific American): https://www.scientificamerican.com/article/the-coronavirus-pandemic-puts-children-at-risk-of-online-sexual-exploitation/
Sex only with those with whom one is self-quarantined	Patient is at risk for infection from sex partner if they have been exposed while outside the home Patient is at risk for infection from an asymptomatic SARS-CoV-2–infected partner
Sex with persons other than those with whom one is self-quarantined	Patient should be counseled on the risk for infection from partners, as well as risk reduction techniques that include minimizing the number of sexual partners, avoiding sex partners with symptoms consistent with SARS-CoV-2, avoiding kissing and sexual behaviors with a risk for fecal–oral transmission or that involve semen or urine, wearing a mask, showering before and after sexual intercourse, and cleaning of the physical space with soap or alcohol wipes COVID-19 and Your Sexual Health (Fenway Health): https://fenwayhealth.org/wp-content/uploads/C19MC-11_Sex-and-COVID-19-Materials_flyer2.pdf Guidance on COVID-19 and sexual health (New York City Department of Health): https://www1.nyc.gov/assets/doh/downloads/pdf/imm/covid-sex-guidance.pdf
Additional resources Building Health Communities Online – Sex Partner Notification Platform: https://tellyourcontacts.org/ What to Know About HIV and COVID-19 (Centers for Disease Control and Prevention) https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/hiv.html COVID-19 Command Center for STD Programs(National Coalition of STD Directors) https://www.ncsddc.org/resource/covid-command-center-for-std-programs/	

COVID-19 = coronavirus disease 2019; SARS-CoV-2 = severe acute respiratory syndrome-coronavirus-2; STD = sexually transmitted disease.

judgment can lead them to withhold information about sexual risk behaviors.

For some patients, complete abstinence from in-person sexual activity is not an achievable goal. In these situations, having sex with persons with whom they are self-quarantining is the safest approach. Those unable to take this approach may benefit from risk reduction counseling (**Table**), which has proven effective in other realms of sexual health (9). Patients should also be provided with information about how to reduce the risk for other sexually transmitted infections as well as the importance of continued use of contraceptives during this time to prevent unwanted pregnancy. The Centers for Disease Control and Prevention have released special guidance regarding SARS-CoV-2 and HIV (10). Those taking HIV preexposure prophylaxis should be encouraged to continue taking this medication consistently (10).

LOOKING TO THE FUTURE

For the foreseeable future, HCPs will need to incorporate new technological advances regarding SARS-CoV-2 into how they think about sexual health and risk. As was seen during the HIV epidemic, antibody tests may play a key role in how we evaluate sexual risk. Though we currently lack data on how long such immunity may last, those who test positive for SARS-CoV-2 antibodies could have relative immunity to the virus. This may allow for the serosorting of individuals for sexual activity, with those testing positive for anti–SARS-CoV-2 antibodies presumed safe to engage in sex together with regard to SARS-CoV-2 transmission, if not for HIV or other sexually transmitted infections. Further research is needed to know if this will be an effective strategy. It will be important for HCPs to proactively discuss with patients what we learn from the emerging science: how reliable the antibody tests are, and to what extent these tests can inform SARS-CoV-2 risk assessment.

As we continue to fight the pandemic, researchers and HCPs ought to keep human sexuality in mind as an important aspect of health and counsel patients whenever possible. Public health officials must continue to disseminate accurate sexual health information. We need to collect more data on the risks related to SARS-CoV-2 transmission through intimate contact, best practices in sexual counseling, and optimal approaches for risk reduction.

From Massachusetts General Hospital and Harvard Medicine School, Boston, Massachusetts (J.L.T.); Massachusetts General Hospital, The Fenway Institute, and Harvard Medicine School, Boston, Massachusetts (A.S.K.); and Beth Israel Deaconess Medical Center, The Fenway Institute, and Harvard Medical School, Boston, Massachusetts (K.H.M.).

Disclosures: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M20-2004.

Corresponding Author: Jack Turban, MD, MHS, Massachusetts General Hospital, 15 Parkman Street, WAC 812, Boston, MA 02114; e-mail, jack.turban@mgh.harvard.edu.

Current author addresses and author contributions are available at Annals.org.

Ann Intern Med. doi:10.7326/M20-2004

References

1. **van Doremalen N, Bushmaker T, Morris DH, et al.** Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1 [Letter]. N Engl J Med. 2020;382:1564-1567. [PMID: 32182409] doi:10.1056/NEJMc2004973

2. **Pan F, Xiao X, Guo J, et al.** No evidence of SARS-CoV-2 in semen of males recovering from COVID-19. Fertil Steril. 2020. doi:10.1016/j.fertnstert.2020.04.024

3. **Qiu L, Liu X, Xiao M, et al.** SARS-CoV-2 is not detectable in the vaginal fluid of women with severe COVID-19 infection. Clin Infect Dis. 2020. [PMID: 32241022] doi:10.1093/cid/ciaa375

4. **Li D, Jin M, Bao P, et al.** Clinical characteristics and results of semen tests among men with coronavirus disease 2019. JAMA Netw Open. 2020;3:e208292. [PMID: 32379329] doi:10.1001/jamanetworkopen.2020.8292

5. **Wang W, Xu Y, Gao R, et al.** Detection of SARS-CoV-2 in different types of clinical specimens. JAMA. 2020. [PMID: 32159775] doi:10.1001/jama.2020.3786

6. **Guan WJ, Zhong NS.** Clinical characteristics of covid-19 in China. reply [Letter]. N Engl J Med. 2020;382:1861-1862. [PMID: 32220206] doi:10.1056/NEJMc2005203

7. **Chen Y, Chen L, Deng Q, et al.** The presence of SARS-CoV-2 RNA in the feces of COVID-19 patients. J Med Virol. 2020. [PMID: 32243607] doi:10.1002/jmv.25825

8. **Ott MA, Santelli JS.** Abstinence and abstinence-only education. Curr Opin Obstet Gynecol. 2007;19:446-52. [PMID: 17885460]

9. **Eaton LA, Kalichman SC, Kalichman MO, et al.** Randomised controlled trial of a sexual risk reduction intervention for STI prevention among men who have sex with men in the USA. Sex Transm Infect. 2018;94:40-45. [PMID: 28404766] doi:10.1136/sextrans-2016-052835

10. **Centers for Disease Control and Prevention.** What to know about HIV and COVID-19. 2020. 18 March 2020. Accessed at www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/hiv.html on 5 May 2020.

Current Author Addresses: Dr. Turban: Jack Turban, MD, MHS, Massachusetts General Hospital, 15 Parkman Street, WAC 812, Boston, MA 02114.
Dr. Keuroghlian: Massachusetts General Hospital, Wang Building, 8th Floor, Boston, MA 02114.
Dr. Mayer: Fenway Health, 7 Haviland Street, Boston, MA 02115.

Author Contributions: Conception and design: J.L. Turban, A. Keuroghlian, K.H. Mayer.
Analysis and interpretation of the data: J.L. Turban.
Drafting of the article: A. Keuroghlian, J.L. Turban.
Critical revision for important intellectual content: J.L. Turban, A. Keuroghlian, K.H. Mayer.
Final approval of the article: J.L. Turban, A. Keuroghlian, K.H. Mayer.
Obtaining of funding: J.L. Turban.
Administrative, technical, or logistic support: A. Keuroghlian.

J Formos Med Assoc. 2004 Aug;103(8):624-8.

The physiological impact of wearing an N95 mask during hemodialysis as a precaution against SARS in patients with end-stage renal disease

Tze-Wah Kao [1], Kuo-Chiang Huang, Yu-Ling Huang, Tun-Jun Tsai, Bor-Shen Hsieh, Ming-Shiou Wu

Affiliations
PMID: 15340662

Abstract

Background and purpose: Most patients with end-stage renal disease (ERSD) visiting our hospital for hemodialysis treatment during the SARS outbreak wore an N95 mask. Data on the physiological stress imposed by the wearing of N95 masks remains limited. This study investigated the physiological impact of wearing an N95 mask during hemodialysis (HD) on patients with ESRD.

Methods: ESRD patients who received regular HD at National Taiwan University Hospital between April to June 2003 were enrolled. Each patient wore a new N95 mask (3M Model 8210) during HD (4 hours). Vital signs, clinical symptoms and arterial blood gas measured before and at the end of HD were compared.

Results: Thirty nine patients (23 men; mean age, 57.2 years) were recruited for participation in the study. Seventy percent of the patients showed a reduction in partial pressure of oxygen (PaO2), and 19% developed various degrees of hypoxemia. Wearing an N95 mask significantly reduced the PaO2 level (101.7 +/- 12.6 to 92.7 +/- 15.8 mm Hg, p = 0.006), increased the respiratory rate (16.8 +/- 2.8 to 18.8 +/- 2.7/min, p < 0.001), and increased the occurrence of chest discomfort (3 to 11 patients, p = 0.014) and respiratory distress (1 to 17 patients, p < 0.001). Baseline PaO2 level was the only significant predictor of the magnitude of PaO2 reduction (p < 0.001).

Conclusion: Wearing an N95 mask for 4 hours during HD significantly reduced PaO2 and increased respiratory adverse effects in ESRD patients.

Related information

MedGen

LinkOut - more resources

Medical
Genetic Alliance
MedlinePlus Health Information

Miscellaneous
NCI CPTAC Assay Portal

Tong et al. Antimicrobial Resistance and Infection Control (2015) 4:48
DOI 10.1186/s13756-015-0086-z

ANTIMICROBIAL RESISTANCE & INFECTION CONTROL

RESEARCH Open Access

Respiratory consequences of N95-type Mask usage in pregnant healthcare workers—a controlled clinical study

Pearl Shuang Ye Tong[1], Anita Sugam Kale[1], Kailyn Ng[1], Amelia Peiwen Loke[1], Mahesh Arjandas Choolani[1], Chin Leong Lim[2], Yiong Huak Chan[3], Yap Seng Chong[1], Paul Anantharajah Tambyah[4] and Eu-Leong Yong[4*]

Abstract

Background: Outbreaks of emerging infectious diseases have led to guidelines recommending the routine use of N95 respirators for healthcare workers, many of whom are women of childbearing age. The respiratory effects of prolonged respirator use on pregnant women are unclear although there has been no definite evidence of harm from past use.

Methods: We conducted a two-phase controlled clinical study on healthy pregnant women between 27 to 32 weeks gestation. In phase I, energy expenditure corresponding to the workload of routine nursing tasks was determined. In phase II, pulmonary function of 20 subjects was measured whilst at rest and exercising to the predetermined workload while breathing ambient air first, then breathing through N95-mask materials.

Results: Exercising at 3 MET while breathing through N95-mask materials reduced mean tidal volume (TV) by 23.0 % (95 % CI −33.5 % to −10.5 %, $p < 0.001$) and lowered minute ventilation (VE) by 25.8 % (95 % CI −34.2 % to −15.8 %, $p < 0.001$), with no significant change in breathing frequency compared to breathing ambient air. Volumes of oxygen consumption (VO_2) and carbon dioxide expired (VCO_2) were also significantly reduced; VO_2 by 13.8 % (95 % CI −24.2 % to −3 %, $p = 0.013$) and VCO_2 by 17.7 %, (95 % CI −28.1 % to −8.6 %, $p = 0.001$). Although no changes in the inspired oxygen and carbon dioxide concentrations were demonstrated, breathing through N95-mask materials during low intensity work (3 MET) reduced expired oxygen concentration by 3.2 % (95 % CI: −4.1 % to −2.2 %, $p < 0.001$), and increased expired carbon dioxide by 8.9 % (95 % CI: 6.9 % to 13.1 %; p <0.001) suggesting an increase in metabolism. There were however no changes in the maternal and fetal heart rates, finger-tip capillary lactate levels and oxygen saturation and rating of perceived exertion at the work intensity investigated.

Conclusions: Breathing through N95 mask materials have been shown to impede gaseous exchange and impose an additional workload on the metabolic system of pregnant healthcare workers, and this needs to be taken into consideration in guidelines for respirator use. The benefits of using N95 mask to prevent serious emerging infectious diseases should be weighed against potential respiratory consequences associated with extended N95 respirator usage.

Trial Registration: The study was registered at clinicaltrials.gov, identifier NCT00265926.

Keywords: N95 respirators, Infection control, Pregnant women, Healthcare workers, Respiratory parameters, Controlled trial

* Correspondence: obghead@nus.edu.sg
PAT and ELY are co-Principal Investigators
[4]Medicine, National University of Singapore, 1E Kent Ridge Road, Level 12, Singapore 119228, Singapore
Full list of author information is available at the end of the article

Background

Lessons learnt on infection control from the Severe Acute Respiratory Syndrome (SARS) pandemic in 2003 have been used to formulate strategies [1, 2] to manage the recent Middle Eastern respiratory syndrome (MERS) [3] and H7N9 influenza outbreaks [4]. These infection control measures include recommendations for increased use of protective filtering face-piece respirators (FFR), such as N95-masks [5] especially during aerosol generating procedures. Existing influenza pandemic control plans in many countries have also incorporated recommendations for more widespread use of FFR [6–8]. When the influenza A H1N1 pandemic was declared in 2009, there were guidelines for universal use of N95-masks despite a lack of scientific evidence for its appropriateness in different health care settings [9, 10]. The N95 FFR has also been recommended for the novel MERS Coronavirus.

Little is known about the effects of N95-masks on the respiratory function of pregnant healthcare workers, who can be subjected to prolonged usage of FFR because of their vulnerability to complications from influenza, varicella, and other pathogens transmitted via the respiratory tract [11]. It is also known that pregnant women have a significantly greater respiratory burden due to factors such as increased oxygen (O_2) demand, increased nasal airway resistance, decreased functional residual capacity due to diaphragmatic splinting; all these contributing to the "physiologic" dyspnea of pregnancy [12]. There are also robust data linking respiratory compromise and adverse perinatal outcomes in women who have chronic respiratory conditions, from large scale studies on women with conditions such as asthma and obstructive sleep apnea. These outcomes include preterm labour, impaired fetal growth, and pre-eclampsia [13, 14].

Balancing the potential benefits of respiratory protection against the possible discomfort [15] and potential additive adverse effects on the respiratory functions of pregnant healthcare workers is difficult in the absence of clear data although there is no definite evidence of harm from decades of use of such respirators [16]. A recent study comparing a cohort of pregnant women between 13 to 35 weeks gestation and non-pregnant women showed no differences in respiratory rate, oxygen saturation and transcutaneous carbon dioxide levels in pregnant compared with non-pregnant subjects wearing the N95 FFR during exercise and sedentary activities for over a 1-hour period [17]. However, that study did not specifically examine the impact on busy healthcare workers. Pregnancy has been reported to be the most common cause for denying medical clearance for N95-mask use in a non-medical setting but the specific adverse effects of the respirator itself have not been documented [18]. Our study was performed to address the limited data on N95-mask usage in pregnancy with the aim of investigating the effects of breathing through the N95 mask materials on respiratory functions at rest, during low intensity work, and recovery thereafter in pregnant healthcare workers. The differences in work of breathing and potential adjustments in respiration that are contributed by pregnancy may provide guidance on the use of N95-masks by pregnant health care workers in high-risk environments.

Methods

This controlled clinical trial was carried out in 2 phases in the Investigational Medicine Unit of National University Hospital (NUH), Singapore. Study procedures were approved by the National Healthcare Group Domain Specific Review Board in July 2010 (Reference Number: 2010/00226), and written informed consent was obtained from all participants.

Healthy women with singleton pregnancies between 27 to 32 weeks gestation were recruited on a voluntary basis from amongst hospital staff and clinic patients. Eight pregnant health care workers were recruited in phases I and 20 pregnant women were recruited in phase II to participate in the study. Subjects were instructed to have adequate rest and to avoid strenuous activity prior to the study to ensure that the tests were conducted under normal lifestyle conditions. All subjects were told to have their meals at least 2 h before start of study. A screening questionnaire was administered to each subject, who then had baseline medical and obstetric examinations prior to participating in the study.

Inclusion and exclusion criteria

Subjects had spontaneously conceived singleton pregnancies and were between 21 to 40 years old. They had no history of cardiorespiratory illness, influenza-like illness in the week prior to the trial, or any pregnancy-related complications such as gestational diabetes, hypertension, intrauterine growth restriction, placenta praevia, ruptured membranes, or threatened preterm labor. They were also free of any neuromuscular conditions that would preclude them from using the treadmill. Their hemoglobin levels were ≥11g/dL, and they did not have any haemoglobinopathies such as thalassemia that could interfere with oxygen carriage in the blood.

Study design

In Phase I, the volume of O_2 uptake (VO_2) corresponding to the workload of routine nursing tasks in the ward was determined. These healthcare workers wore, and breathed through, a tight-fitted respiratory mask (Hans Rudolph, V-mask, Kansas) that was attached by a harness to a portable telemetric metabolic cart [19–21] (K4b2, Cosmed s.r.l, Rome, Italy) while moving about freely performing simulated routine nursing tasks in a specific order, such as walking around the ward, sponging

and transferring mannequins from beds to chairs with another assistant (Fig. 1). Their average work intensity was determined with VO_2 (ml/kg/min) measurements and converted to the corresponding metabolic equivalents (MET) to gauge the energy expenditure (1 MET is equal to 3.5 ml/kg/min of O_2 consumed).

In Phase II, the respiratory effects of wearing N95 masks were examined. Each subject underwent two 15-minute exercise cycles on a treadmill. Each subject wore a Hans Rudolph mask, similar to that in Phase I, attached to a laboratory-based metabolic cart (Cortex Metalyser 3BR2, Leipzig, Germany) in order to obtain real time respiratory parameters during exercise. In the first (control) cycle, subjects wore the Hans Rudolph mask with the outlet opened to ambient air. In the second (N95) cycle, outlets of the Hans Rudolph masks were covered by materials obtained from representative supplies of N95 masks (3M, St. Paul, MN, USA). N95-mask materials were trimmed to form an airtight seal over the Hans Rudolph mask outlet so that the air flow resistance on inspiration and expiration would come from the mask material, simulating the actual wearing of an N95 respirator (Fig. 2) This experimental design allowed each subject to act as her own control.

The fine adjustments were made to the treadmill speed every 3 min to maintain energy expenditure at 3 MET. Similar treadmill speed profile was repeated in the second exercise cycle for each subject. For both Control and N95 cycles, respiratory parameters were measured during an initial 10-minute rest period, followed by a 15-minute exercise period, and subsequently a 25-minute rest period. There was a 30-minute break between control and N95 cycles. Prior to the N95 cycle, an additional 15-minute conditioning period was allowed to enable patient to adapt

Fig. 1 Determination of average work intensity of Health care workers: In phase I pregnant subjects performed simulated patient care activities while breathing through a tight fitting mask with a pneumotachometer. Oxygen content was sampled at every breath and measured with a portable telemetric metabolic cart

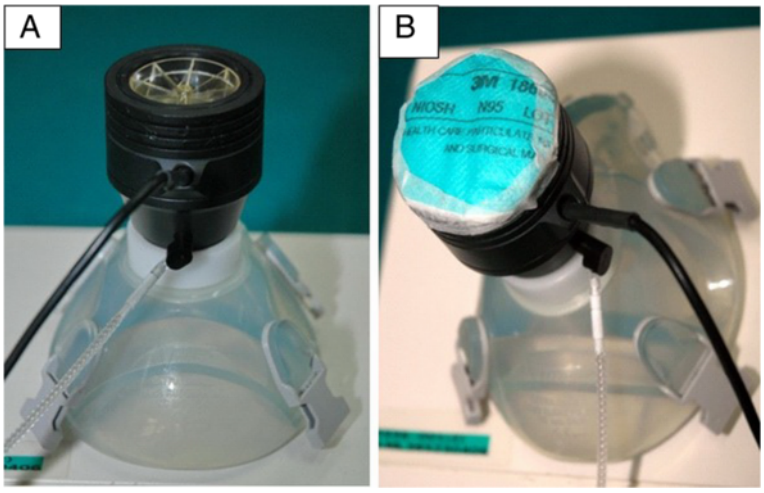


Fig. 2 Tight fitting Hans Rudolph respirator masks used in Phase II. (**a**) Control cycles with outlet open to air, and (**b**) N95 cycles with outlet covered by N95 mask materials

to N95-respirator conditions. The subjects breathed through the N95 mask materials continuously throughout the N95 cycles (Fig. 3).

A cardiotocography (CTG) was performed prior to the study and after each exercise-cycle. Finger-prick lactate concentrations were measured immediately before and after each exercise-cycle with Lactate Pro (Arkray Global Business Inc). The Borg Scale questionnaire [22] was administered after each exercise cycle to measure the rating of perceived exertion. The Borg Scale ranges from 6 for "no feeling of exertion," to 20 which corresponds to "very, very hard exertion". A fully equipped resuscitation cart was present throughout the study with trained medical personnel available to immediately address any medical concerns of the subjects.

Equipment for measurement of pulmonary function and its calibration

In both phases, the participants wore a tight fitting mask (Hans Rudolph) that was attached to the metabolic cart

PHASE I PROTOCOL (n= 8)

Determination of average workload of nurses by performing simulated patient care activities while breathing through a tight fitting mask with pneumotachometer. Oxygen content was sampled at every breath and measured with portable telemetric metabolic cart.

PHASE II PROTOCOL (n= 20)

Control cycle
(wearing mask with outlet open to air)

Period 1 (rest): 10 min
Respiratory parameters measured while resting
Period 2 (exercise): 15 min
Respiratory parameters measured while exercising on treadmill with speed set to the predetermined work load
Period 3 (rest): 25 min
Respiratory parameters measured while resting after exercise

Break of 30 mins (without mask)

N95 cycle
(wearing mask with outlet covered by N95 mask materials)

15 min conditioning period with mask at rest
Period 1 (rest): 10 min
Respiratory parameters measured while resting
Period 2 (exercise): 15 min
Respiratory parameters measured while exercising on treadmill with speed set to the predetermined work load
Period 3 (rest): 25 min
Respiratory parameters measured while resting after exercise

Fig. 3 Phase I & 2 Protocols

through an air sampling tube. Inspired ambient air and expired air were channeled through a pneumotachometer that was attached to the front of the mask which calculated air volume by the rate of rotation of a rotor turbine located within it. The turbine had zero resistance to air flow and the rate of rotation of the turbine, sensed by infrared light within the pneumotachometer, corresponds directly to inspired and expired air volume for each breath. Multiple air samples from each expired-breath was drawn into the metabolic carts through a sampling line for the measurement of oxygen and carbon dioxide content by the respective gas sensors within the metabolic carts. From this data, the following parameters were calculated: volumes of oxygen (VO_2) and carbon dioxide (VCO_2) exchanged, breathing frequency (BF), tidal volume (TV), minute ventilation (VE), forced expired O_2 (FeO_2), forced expired CO_2 ($FeCO_2$), forced inspired end-tidal O_2 and CO_2 concentrations ($Fi_{et}O_2$, $Fi_{et}CO_2$).

The calibration procedures for both portable and laboratory-based metabolic carts were carried out according to the manufacturer's instructions and were performed daily to ensure uniformity in their measurements. The study was conducted in a standardized air-conditioned room similar to a hospital ward with constant humidity and temperature.

Criteria for study termination

The study was to be terminated if: (1) the CTG revealed that the fetus was adversely affected by the mother's activity on the treadmill either by a suspicious or pathological trace as defined by the National Institute for Health and Care Excellence [23],(2) the subject was unable to complete the trial for any reason including breathlessness or pain, (3) injury was sustained as a result of the exercise, or (4) maternal heart rate was >155 beats/min [24, 25].

Statistical analysis

Since there were many respiratory variables of interest, the sample size calculations were performed on an overall picture that there was at least a 20 % difference between breathing through N95 mask materials vs control for any of respiratory variable of interest. Postulating that breathing through N95 mask materials would have an at least 20 % variation (with standard deviation 25 %) from the control respiratory variables, recruitment of 20 subjects would have a 90 % power and a 2-sided p-value of 5 % to show a statistically significant result. A mixed linear model analysis (to handle paired observations) adjusting for relevant covariates was performed. All analyses were performed using IBM SPSS version 20.0 (Armonk, NY) and statistical significance was set at $p < 0.05$.

Results

Twenty-eight pregnant women were recruited between September 2010 and September 2011. In Phase I, the mean and SEM of VO_2 was 9.04 (±0.75) ml/kg/min, which was equivalent to about ~3 MET. All 8 subjects enrolled in Phase I fulfilled the inclusion criteria and completed the study. Subjects in Phase II were then subjected to this workload. For Phase II, 23 subjects were screened. Two subjects were excluded because of maternal anemia, and a third was excluded because the fetus exhibited ventricular bigeminy on ultrasound examination prior to starting the study. One subject experienced uterine contractions and did not complete the study, bringing the total number of subjects who completed Phase II of the study to 19. There were no other trial terminations due to adverse events.

The mean age of the 19 subjects in Phase II was 30.0 (±0.87) years, their average gestation was 30.1(±0.28) weeks, and mean BMI was 26.6 kg/m^2 (±1.4). Of the cases, 10 were primigravidas and 9 were multigravidas. There were 13 nurses, 6 homemakers, and 9 women doing administrative work.

Effect on tidal volumes, breathing frequency, minute ventilation

During the pre-exercise rest period, breathing through N95-mask materials lowered TV by a mean of 0.15 L compared with controls (95 % CI: −0.23, −0.08; $p < 0.001$) (Fig. 4a). TV during both exercise cycles increased rapidly to reach a plateau which was about 50 % higher than that observed at rest, within a minute. However, compared with controls, exercising with N95-masks reduced mean TV by 0.21L (95 % CI: −0.32, −0.10; $p < 0.001$), a 23 % decrease (Fig. 4a, Table 1). Mean BF increased by 35 % during exercise for both control and N95 cycles compared to the rest period, but there was no difference in BF with and without wearing N95-masks (Fig. 4b, Table 1). Wearing N95-respirators lowered V_E by 25.8 %, a mean difference of 5.55L/min (95 % CI: −7.58, −3.51; $p < 0.001$) (Fig. 4c, Table 1). Significant differences in TV and V_E with N95 cycles persisted in the post-exercise rest period (Table 1). There was a tidal volume reduction of 0.08L ($p = 0.02$) and minute ventilation reduction of 1.1L/min ($p = 0.031$).

Effect on O_2 and CO_2 concentrations in inspired and expired air

Forced expired O_2 concentration (FeO_2) during the pre-exercise rest period decreased by 0.52 % (95 % CI: −0.79, −0.25; $p = 0.001$) with N95-mask use versus controls (Fig. 5a, Table 1). Wearing of N95-masks during exercise reduced FeO_2 by 0.54 % compared with controls (95 % CI: −0.70, −0.38; $p < 0.001$). Reduction in FeO_2 with the use of N95-masks persisted in the post-exercise rest period. Concomitantly, forced

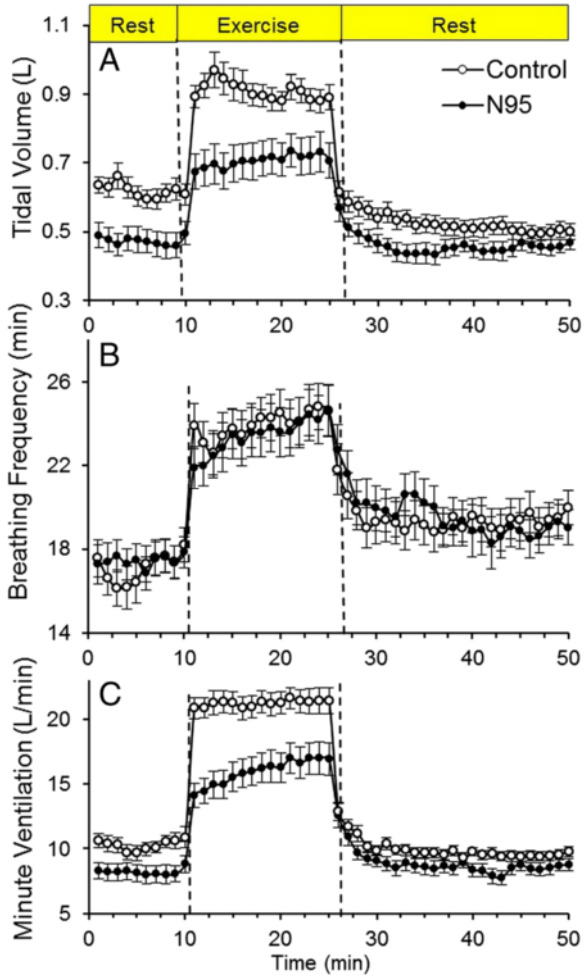


Fig. 4 Ventilation functions in pregnancy during rest and exercise. Inspired and expired air volumes were measured with a pneumotachometer attached to mask outlet to obtain (**a**) tidal volume, (**b**) breathing frequency and (**c**) minute ventilation. $N = 19$ (±SEM)

expired CO_2 concentration ($FeCO_2$) was significantly elevated with N95-mask use in the pre-exercise, exercise and post exercise periods compared with no mask usage (Fig. 5b, Table 1). During the exercise period, the wearing of N95-masks resulted in increase in $FeCO_2$ of 0.30 % (95 % CI: 0.18, 0.42; $p < 0.001$) compared with controls. (Fig. 5b, Table 1). In contrast, no significant differences were observed in the inspired oxygen (FiO_2) or carbon dioxide ($FiCO_2$) concentrations of inspired air before, during or after exercise (Fig. 6, Table 1).

Effect on pulmonary gas exchange

When performing work on the treadmill equivalent to 3 MET, VO_2 and VCO_2 increased by about two-fold for all subjects compared to the rest periods. Strikingly, wearing of the N95-mask during exercise resulted in lowering VO_2 by 13.8 %, a mean of 1.30 ml/min/kg (95 % CI: −2.30, −0.31; $p = 0.013$) (Fig. 7a, Table 1). Similarly, VCO_2 was lowered by 17.7 %, a mean of 0.10 ml/min/kg (95 % CI: −0.15, −0.05; $p = 0.001$) (Fig. 7b, Table 1).

Table 1 Changes in respiratory parameters of pregnant subjects breathing through N95 masks compared to controls breathing ambient air

N95 Mask vs Control	Mean difference	SE (95 % CI)	*P* value
Pre exercise rest period			
VO_2	−0.40	0.30 (−1.02, 0.23)	0.20
VCO_2	−0.04	0.02 (−0.07, −0.003)	**0.035**
Tidal volume	−0.15	0.04 (−0.23, −0.08)	**< 0.001**
Breathing frequency	0.31	0.69 (−1.13, 1.75)	0.66
Minute ventilation	−2.23	0.66 (−3.60, −0.85)	**0.003**
FeO_2	−0.52	0.97 (−0.79, −0.25)	**0.001**
$FeCO_2$	0.25	0.07 (0.11, 0.40)	**0.002**
Fi O_2	−0.02	0.07 (−0.16, 0.12)	0.76
Fi CO_2	−0.01	0.01 (−0.03, 0.01)	0.47
Exercise period			
VO_2	−1.30	0.47 (−2.3, −0.31)	**0.01**
VCO_2	−0.10	0.25 (−0.15, −0.05)	**0.001**
Tidal volume	−0.21	0.05 (−0.32, −0.11)	**< 0.001**
Breathing frequency	−0.51	0.67 (−1.92, 0.89)	0.45
Minute ventilation	−5.55	0.97 (−7.58, −3.51)	**< 0.001**
FeO_2	−0.54	0.08 (−0.70, −0.38)	**< 0.001**
$FeCO_2$	0.30	0.06 (0.18, 0.42)	**< 0.001**
Fi O_2	0.02	0.08 (−0.13, 0.17)	0.81
Fi CO_2	0.004	0.01 (−0.23, 0.03)	0.75
Post exercise rest period			
VO_2	−0.22	0.21 (−0.67, 0.23)	0.32
VCO_2	−0.01	0.01 (−0.04, 0.01)	0.29
Tidal volume	−0.08	0.03 (−0.14; −0.01)	**0.02**
Breathing frequency	0.48	0.50 (−0.56, 1.53)	0.34
Minute ventilation	−1.10	0.47 (−2.10, −0.11)	**0.031**
FeO_2	−0.30	0.09 (−0.49, −0.11)	**0.004**
$FeCO_2$	0.19	0.05 (0.09, 0.29)	**0.001**
Fi O_2	−0.03	0.09 (−0.2, 0.16)	0.77
Fi CO_2	0.00	0.02 (−0.03, 0.03)	0.99

(bolded values: statistically significant)

Effect on maternal and fetal physiological parameters

For all subjects, overall maternal heart rate increased from 89 ± 1.8 to 107 ± 1.9 beats/min with exercise. There were no significant difference in heart rate between the N95-masks and control cycles. There were also no changes in basal fetal heart rates (mean heart rate of 133 beats per minute) or variability (15–16 beats per minute) in all the CTGs. There were no significant differences in

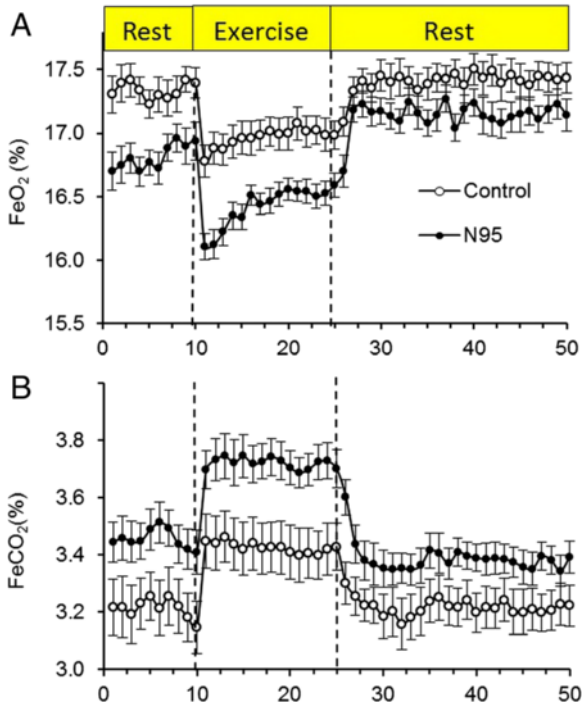


Fig. 5 Oxygen and Carbon dioxide content of expired air in pregnant women at rest and following exercise. (**a**) Forced expired O_2 (FeO_2) and (**b**) forced expired CO_2 concentrations ($FeCO_2$) were measured with a metabolic cart. $N = 19$ (±SEM)

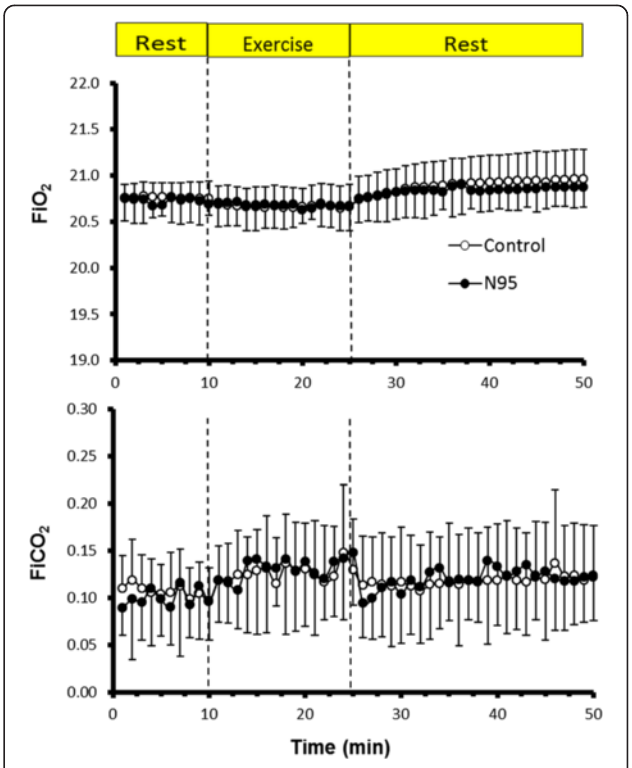


Fig. 6 Oxygen and Carbon dioxide content of inspired air in pregnant women at rest and following exercise. (**a**) Forced inspired O_2 (FiO_2) and (**b**) forced inspired CO_2 concentrations ($FiCO_2$) were measured with a metabolic cart. $N = 19$ (±SEM)

lactate levels pre exercise (1.8 ± 0.2 mmol/L), post exercise breathing through ambient air (1.6 ± 0.2 mmol/L), and post exercise with the N95-mask (2.1 ± 0.4 mmol/L). There were no differences in finger-tip capillary oxygen saturation levels with the mask and without the mask; 98.3 ± 0.18 % and 98.4 ± 0.11 % respectively. The Borg scale indicated that exercise induced a borderline increase in perceived effort from being 9.1(±0.60) to 10.7(±0.8) after the N95 cycles. These parameters did not reach statistical significance (Table 2).

Discussion

We found that in women in mid-pregnancy, breathing through the N95 respirator material when performing low intensity work significantly reduced VO_2 (13.8 %) and VCO_2 (17.7 %), which was due to a corresponding decrease in V_E (25.8 %) and TV (23 %), without a compensatory increase in BF. This decrease in air intake volume, together with unchanged concentrations of inspired O_2 and CO_2 imply a decrease in overall amount of O_2 and CO_2 inspired. Coupled with a 3.2 % decrease in FeO_2 and an 8.9 % increase in $FeCO_2$, these results suggest an increased consumption of O_2 and production of CO_2, leading to possible concerns regarding prolonged usage of N95-masks on respiratory functions in pregnant women performing physical work. The decrease in V_E, TV, FeO_2

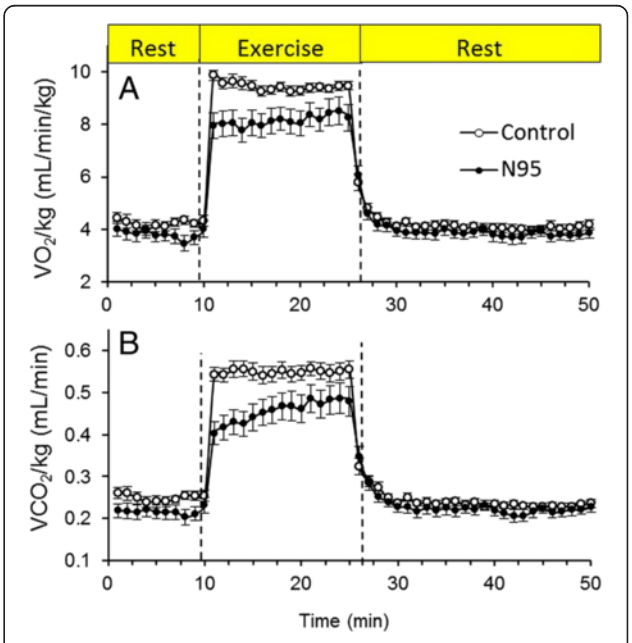


Fig. 7 Pulmonary gas exchange in pregnant women at rest and following exercise. O_2 and CO_2 concentrations were measured with a metabolic cart to calculate volumes of oxygen (**a**) VO_2, or carbon dioxide (**b**) VCO_2 exchanged with each breath. $N = 19$ (±SEM)

Table 2 Changes in maternal and fetal physiological parameters breathing through N95 masks compared to controls breathing ambient air

N95 Mask vs Control	Mean difference	SE (95 % CI)	*P* value
Pre exercise maternal heart rate	−2.0	1.3 (−4.7; 0.6)	0.11
Exercise maternal heart rate	−0.5	1.4 (−3.4; 2.4)	0.71
Post exercise maternal heart rate	−0.8	1.6 (−4.1; 2.5)	0.61
Baseline fetal heart rate	−0.5	0.6 −1.6; 0.6)	0.33
Capillary lactate	0.5	0.5 (−0.3; 1.4)	0.17
Finger-tip capillary oxygen saturation	0.1	0.02 (−0.06; 0.14	0.32
Borg scale	1.6	0.9 (−0.01; 3.3)	0.060

and the increase in $FeCO_2$ were also significant during the rest periods. These results suggest that breathing through the N95 mask material can limit the overall volume of amount of oxygen intake and also increase the rate of metabolism.

When performing work equivalent to routine bedside nursing with N95-masks, non-pregnant subjects have been previously reported to maintain their V_E compared with controls, with non-significant changes in the TV and BF [26]. Other studies in both non-pregnant and pregnant subjects have shown, on the contrary, an increase or decrease in BF but the TV and V_E were not measured in these trials [17, 27, 28]. In contrast, our pregnant subjects were unable to proportionately increase both TV and BF to maintain their V_E during rest and in response to exercise while breathing through N95-masks materials. A significant 26 % reduction in $V_{E,}$ with a 23 % reduction in mean TV was noted during exercise, possibly due to diaphragmatic splinting. This decrease in V_E led to a corresponding decrease in VO_2 (13.8 %) and VCO_2 (17.7 %) (Fig. 7). The decrease in FeO_2 and increase in $FeCO_2$ are likely due to a stimulation of the respiratory drive resulting in greater efforts required to breathe through the N95 mask materials and a concomitant greater extraction of O_2 for aerobic metabolism. These results also suggest that the N95 mask may impede gaseous exchange resulting in hypoventilation. Although the subjects appeared to adapt to the increased workload with no changes in the other maternal and fetal physiological parameters with no evidence of hypoxia found on finger-tip capillary oxygen saturation nor increased lactic acid production to suggest a shift toward anaerobic respiration, our results suggest that performing physical work with the N95-mask appears to stress the aerobic metabolism and increase the CO_2 load within the circulation.

In non-pregnant subjects, it has been shown that use of N95-respirators can increase CO_2 levels within the masks by 1.8–3 %, suggesting that the increase in expired CO_2 concentration could also be due to the accumulation of expired CO_2 trapped in the dead space of the N95 mask [29, 30]. Our results do not support such a view because $FiCO_2$ did not increase even with the use of N95 materials and total CO_2 intake was reduced due to the corresponding decrease in V_E. These results further affirm that the increase in expired CO_2 mainly arose from increased rate of aerobic metabolism. The higher circulating CO_2 concentration observed in our study was in agreement with increase in transcutaneous CO_2 observed in pregnant and non-pregnant women after a 20-minute exercise wearing a respirator as compared to not wearing the respirator [17]. It must be borne in mind that CO_2 levels in the blood of pregnant women are usually lower due to physiological hyperventilation. An increase in forced expired CO_2 which reflects a rise in blood CO_2 levels hence contributes to impair elimination of fetal CO_2 as arterial CO_2 is normally reduced in pregnancy to allow for a steeper diffusion gradient of CO_2 from fetal blood across to the mother.

Our study was limited in that we were unable to evaluate the effects of N95-mask usage at higher work intensities and over longer durations because of ethical concerns. To ascertain safety, parameters such as maternal and fetal heart rate changes, lactate and finger-tip capillary oxygen saturation levels were monitored to ensure that no significant hypoxemia was induced in our subjects and their fetuses.

Wearing the Hans Rudolph mask with aperture occluded by the mask material allows accurate measurement of respiratory parameters despite its use not being exactly the same as the standard use of the N95-mask. However this is the closest way of obtaining accurate physiologic data to best characterize the impact of the materials used in the respirators. The other limitation of our study was the narrow window of gestation studied (between 27 to 32 weeks). These group of women were deemed to be representative of pregnant women as they would have undergone marked respiratory adaptations to pregnancy, especially diaphragmatic splinting from the enlarging uterus. However, it is postulated that, in situations of exercising at a higher intensity, prolonged N95-mask usage or at more advanced gestations, a

greater degree of oxygen deficit due to a corresponding decrease in V_E can have more marked effects on the respiratory function of pregnant women. We also focused exclusively on pregnant women without using non-pregnant controls to better define the risks in this group which are the most controversial.

Our study demonstrates, for the first time, that pregnant women in mid-pregnancy are unable to maintain their minute ventilation while breathing through N95-mask materials. There is also a decrease oxygen uptake and increase carbon dioxide production as a result of the increased workload on breathing imposed by mask use, both at rest and low work intensity. This supports the view of some that pregnant healthcare workers should probably refrain from prolonged N95-mask use towards the third trimester. The complications to the women and her fetus that can result from a prolonged decrease in V_E and increased work of breathing, are unknown.

Conclusions

While there is a substantial negative change in TV, V_E and the VO_2 and VCO_2 exchanged, there was no impact of breathing through N95 mask materials, to finger-tip oxygen saturation, maternal or fetal heart rate and no drive to increase BF in pregnancy compared to breathing ambient air at the level of exercise in our study. This study shows important descriptive findings of changes to respiratory physiology with mask use, which do not appear to have sufficient significant clinical impact based on the parameters monitored, that had been deliberately kept within the normal ranges to ensure safety of the subjects. Although harm was not demonstrated in the context of this experimental protocol, the significant changes to respiratory physiology caused by breathing through N95 mask materials raise the concern regarding prolonged use of N95-masks by pregnant healthcare workers. Our results suggest that pregnant women may experience more fatigue and require more rest breaks from mask use. Scheduled work breaks should be considered for pregnant healthcare workers working in high risk areas which require prolonged use of N95 respirators. In face of the imminent threat of pandemic airborne respiratory diseases it should be emphasized that the benefit of using N95 mask to prevent serious emerging infectious diseases should be weighed against possible respiratory consequences associated with N95 mask usage. Use of alternative protective methods such as surgical masks with lesser airway resistance, should be considered in appropriate settings [31–33]. These have been shown to be equally effective for the prevention of droplet infections such as influenza [31], although they are insufficient for protection against airborne pathogens. Innovative interventions to improve the design of FFR, are urgently needed in the face of the imminent threat of pandemics of acute airborne respiratory infections. The key is to ensure healthcare worker protection from infectious agents without jeopardizing the wellbeing of the pregnant healthcare workers and their fetuses.

Competing interest

The authors declare that they have no competing interests.

Authors' contributions

PST, MAC, YSC, CLL, PAT, ELY designed the trial. PST, APL, KN, ASK, CLL, YSC, ELY were responsible for recruitment of subjects, organisation and conduct of the trial. YHC analysed the data and advised on statistical issues. PST wrote the first draft of the report. MAC, YSC, CLL, PAT, ELY contributed to redrafting. All authors read and approved the final manuscript.

Acknowledgements

We would like to extend our gratitude to the Defence Science Organisation for their technical support during the study, and to the Department of Obstetrics and Gynaecology, National University Health System for funding the study.

Author details

[1]Department of Obstetrics and Gynecology, National University Hospital, 11 Mandalay Road, Singapore 308232, Singapore. [2]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, 11 Mandalay Road, Singapore 308232, Singapore. [3]Biostatistics Unit, National University of Singapore, Singapore, Republic of Singapore. [4]Medicine, National University of Singapore, 1E Kent Ridge Road, Level 12, Singapore 119228, Singapore.

Received: 15 July 2015 Accepted: 22 October 2015
Published online: 16 November 2015

References

1. Tambyah PA. Severe acute respiratory syndrome from the trenches, at a Singapore university hospital. Lancet Infect Dis. 2004;4:690–6.
2. Centers for Disease Control and Prevention, (15 May 2014). Interim Infection Prevention and Control Recommendations for Hospitalised Patients with Middle East Respiratory Syndrome Coronavirus (MERS-CoV). http://www.cdc.gov/coronavirus/mers/infection-prevention-control.html. Accessed 28 October 2015.
3. Pollack MP, Pringle C, Madoff LC, Memish ZA. Latest outbreak news from ProMED-mail: novel coronavirus – Middle East. Int J Infect Dis. 2013;17:e143–4.
4. Gao R, Cao B, Hu Y, Feng Z, Wang D, Hu W, et al. Human infection with a novel avian-origin Influenza A (H7N9) virus. N Engl J Med. 2013;368:1888–97.
5. Seto WH, Tsang D, Yung RW, Ching TY, Ng TK, Ho M, et al. Advisors of Expert SARS group of Hospital Authority. Effectiveness of precautions against droplets and contact in prevention of nosocomial transmission of severe acute respiratory syndrome (SARS). Lancet. 2003;361:1519–20.
6. Homeland Security Council, (July 2007). National Strategy for Pandemic Influenza: Implementation Plan. http://www.flu.gov/planning-preparedness/federal/pandemic-influenza-oneyear.pdf. Accessed 28 October 2015.
7. Ministry of Health Singapore, (April 2014). Pandemic readiness and response plan for influenza and other acute respiratory disease. http://www.moh.gov.sg/content/dam/moh_web/Diseases%20and%20Conditions/DORSCON%202013/Interim%20Pandemic%20Plan%20Public%20Ver%20_April%202014.pdf. Accessed 28 October 2015.
8. The Government of the Hong Kong Special Administrative Region, (2014). Preparedness Plan for Influenza Pandemic. http://www.chp.gov.hk/files/pdf/erib_preparedness_plan_for_influenza_pandemic_2014_eng.pdf. Accessed 28 October 2015.
9. Radonovich Jr LJ, Perl TM, Davey V, Cohen H. Preventing the soldiers of health care from becoming victims on the pandemic battlefield: respirators or surgical masks as the armor of choice. Disaster Med Public Health Prep. 2009;3 Suppl 2:S203–10.
10. Jefferson T, Del Mar C, Dooley L, Ferroni E, Al-Ansary LA, Bawazeer GA, et al. Physical interventions to interrupt or reduce the spread of respiratory viruses: systematic review. BMJ. 2009;339:b3675.
11. Jamieson DJ, Honein MA, Rasmussen SA, Williams JL, Swerdlow DL, Biggerstaff MS, et al. H1N1 2009 influenza virus infection during pregnancy in the USA. Lancet. 2009;374:451–8.

12. Bobrowski RA. Pulmonary physiology in pregnancy. Clin Obstet Gynecol. 2010;53:285–300.
13. Murphy VE, Namazy JA, Powell H, Schatz M, Chambers C, Attia J, et al. A meta-analysis of adverse perinatal outcomes in women with asthma. BJOG. 2011;118:1314–23.
14. Fung AM, Wilson DL, Barnes M, Walker SP. Obstructive sleep apnea and pregnancy: the effect on perinatal outcomes. J Perinatol. 2012;32:399–406.
15. Lim EC, Seet RC, Lee KH, Wilder-Smith EP, Chuah BY, Ong BK. Headaches and the N95 face-mask amongst healthcare providers. Acta Neurol Scand. 2006;113:199–202.
16. Roberge RJ. Physiological burden associated with the use of filtering facepiece respirators (N95 masks) during pregnancy. J Womens Health (Larchmt). 2009;18:819–26.
17. Roberge RJ, Kim JH, Powell JB. N95 respirator use during advanced pregnancy. Am J Infect Control. 2014;42:1097–100.
18. Pappas GP, Takaro TK, Stover B, Beaudet N, Salazar M, Calcagni J, et al. Respiratory protective devices: rates of medical clearance and causes for work restrictions. Am J Ind Med. 1999;35:390–4.
19. Duffield R, Dawson B, Pinnington HC, Wong P. Accuracy and reliability of a Cosmed K4b2 portable gas analysis system. J Sci Med Sport. 2004;7:11–22.
20. Leprêtre PM, Weissland T, Paton C, Jeanne M, Delannaud S, Ahmaidi S. Comparison of 2 Portable Respiratory Gas Analysers. Int J Sports Med. 2012;33:728–33.
21. McLaughlin JE, King GA, Howley ET, Bassett Jr DR, Ainsworth BE. Validation of the COSMED K4 b2 portable metabolic system. Int J Sports Med. 2001;22:280–4.
22. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc. 1982;14:377–81.
23. NICE guidelines [CG 55] September 2007. Intrapartum care: Care of healthy women and their babies during childbirth. http://www.nice.org.uk/guidance/cg190. Accessed 28 October 2015.
24. Artal R, O'Toole M. Guidelines of the American College of Obstetrics and Gynecology for exercise during pregnancy and the postpartum period. British Journal of Sports Medicine. 2004;37:6–12.
25. Lokey EA, Tran ZV, Wells CL, Myers BC, Tran AC. Effects of physical exercise on pregnancy outcomes: a meta-analysis review. Medicine and Science in Sport and Exercise. 1991;23:1234–9.
26. Roberge RJ, Coca A, Williams WJ, Powell JB, Palmiero AJ. Physiological impact of the N95 filtering facepiece respirator on healthcare workers. Respir Care. 2010;55:569–77.
27. Jones JG. The physiological cost of wearing a disposable respirator. Am Ind Hyg Assoc J. 1991;52:219–25.
28. Kim JH, Benson SM, Roberge RJ. Pulmonary and heart rate responses to wearing N95 filtering facepiece respirators. Am J Infect Control. 2013;41:24–7.
29. Laferty EA, McKay RT. Physiologic effects and measurement of carbon dioxide and oxygen levels during qualitative respirator fit testing. J Chem Health Safety. 2006;13:22–8.
30. Tang JW, HC Willem, TM Ng, Kwok WT. Short-term measures of carbon dioxide levels, physiological indicators and subjective comfort of healthcare workers wearing N95 masks. Influenza Other Respir Viruses. 2011;5(Supp s1):365-66.
31. Loeb M, Dafoe N, Mahony J, John M, Sarabia A, Glavin V. Surgical mask vs N95 respirator for preventing influenza among health care workers: a randomized trial. JAMA. 2009;302:1865–71.
32. Gralton J, McLaws ML. Protecting healthcare workers from pandemic influenza: N95 or surgical masks? Crit Care Med. 2010;38:657–67.
33. Diaz KT, Smaldone GC. Quantifying exposure risk: surgical masks and respirators. Am J Infect Control. 2010;38:501–8.

Dr Ranasinghe). In addition, when performing direct laryngoscopy in intubated patients, many anaesthetists place the laryngoscope blade behind the tube rather than in the vallecula and so lift the epiglottis directly.

It may be that the view achieved in both these situations is more akin to that revealed using a straight bladed laryngoscope – which pulls the epiglottis superiourly using traction exerted on its dorsal surface. Alternatively, the presence of a tracheal tube may obstruct the line of sight, impeding laryngoscopy. Overestimating difficulty is less hazardous than being falsely reassured of an easy conventional intubation.

If one is concerned about possible difficult airway management after extubation, it is safest to remove the tube over a bougie or airway exchange device such as an Aintree Catheter. While conventional laryngoscopy following an awake fibreoptic intubation *may* provide useful information about a patient's airway, making decisions about the degree of difficulty that can be expected after extubation may be unwise, even ignoring the question of tumour growth, fibrosis, oedema progression, etc. The specificity of this manoeuvre, as defined by Yentis (proportion of 'easy' intubation patients who are correctly identified) [2] is, like all other predictors, less than unity.

The goal of airway management is to minimise the number of failed intubations/ventilation, not to minimise the number of awake intubations. Perhaps a useful maxim may be that once a patient has been fibreoptically intubated awake, one needs an unimpeachable reason not to do so again in future.

A. Higgs
O. Hargrove
Warrington Hospital
Warrington WA5 1QG, UK
E-mail: andyhiggs@doctors.org.uk

References

1 Ranasinghe DN. Assessing the laryngeal view by conventional laryngoscopy after awake fibreoptic tracheal intubation. *Anaesthesia* 2006; **61**: 517.
2 Yentis SM. Predicting difficult intubation – worthless exercise or pointless ritual? *Anaesthesia* 2002; **57**: 105–9.

A reply

I would like to thank Drs Higgs and Hargrove for their letter, which I read with interest. They make several valid and important points regarding management of the airway. I agree with their views.

D. N. Ranasinghe
Colchester General Hospital
Colchester CO4 5JL, UK
E-mail: David.Ranasinghe@essexrivers.nhs.uk

Carbon dioxide re-breathing with close fitting face respirator masks

Guidelines for the use of personal protective equipment when caring for patients who are affected by highly infectious pathogens such as SARS, avian influenza and tuberculosis include the use of high efficiency respirators that filter at least 95% of particles with a median diameter of 0.1 micron. These must therefore be close fitting to prevent air leakage [1, 2]. A healthy intensivist (SF), wearing such a respirator (Tecnol Fluidshield PFR95, Kimberly Clark Corporation, Roswell, GA) to perform a percutaneous tracheostomy on a patient with multidrug resistant pulmonary tuberculosis, experienced dyspnoea, tachycardia and tremor after 30 min. End-tidal carbon dioxide measured at the mouth by hand-held capnometry was 6.3 kPa (normal value 5.3 kPa). We postulated that the symptoms were due to hypercapnia.

We measured the end-tidal carbon dioxide levels in four anaesthetists wearing the same design of mask, before and after performing tracheal intubation on another patient with pulmonary tuberculosis. Measurements were taken by sidestream capnometer (Poet LT, Criticare, Waukesha, WI) using a 15-mm T-piece held between the lips. The mean baseline end-tidal carbon dioxide level was 5.18 kPa. Post-procedure (20 min later) the mean end-tidal carbon dioxide level was 5.95 kPa ($p = 0.007$). No subjects reported symptoms of hypercapnia. The rise in end-tidal carbon dioxide is due to rebreathing of expired alveolar gas that is 'trapped' in the respirator, with the degree of rebreathing being proportional to the volume of the respirator ('dead space'). It is likely that all tight-fitting, high efficiency respirators will behave similarly, with only the size of dead space varying between designs.

The respiratory response to hypercapnia is an increase in minute ventilation, giving rise to the sensation of dyspnoea. Moderate (6.18 kPa) to high (7.5 kPa) levels of end-tidal carbon dioxide have also been shown to impair significantly cognitive and psychomotor performance and it is likely that this effect of carbon dioxide is dose related with no threshold [3]. Clearly, our findings are of uncertain practical significance and further trials would be required employing cognitive and psychomotor measurements and arterial blood gas analysis.

In the event of an influenza pandemic, large numbers of healthcare workers may need to wear these respirators for prolonged periods and problems with hypercapnia might reduce the tolerability of these devices. Whether psychomotor performance is affected also remains to be seen.

S. J. Fletcher
M. Clark
P. J. Stanley
Bradford Teaching Hospitals
Bradford BD9 6RJ, UK.
E-mail: sjfletcher@doctors.org.uk.

References

1 The Interdepartmental Working Group on Tuberculosis. *The Prevention and Control of Tuberculosis in the United Kingdom: UK Guidance on the Prevention and Control of Transmission of 1. HIV-related Tuberculosis, 2. Drug-resistant, Including Multiple Drug Resistant, Tuberculosis*. London: Department of Health, The Scottish Office, The Welsh Office, 1998.
2 World Health Organization. *Influenza A (H5N1): WHO Interim Infection Control Guidelines for Health Care Facilities*. Geneva: WHO, 2004.
3 Fothergill DM, Hedges D, Morrison JB. Effects of CO_2 and N_2 partial pressures on cognitive and psychomotor performance. *Undersea Biomedical Research* 1991; **18**: 1–19.

Ergonomics. 2013;56(5):781-90. doi: 10.1080/00140139.2013.777128. Epub 2013 Mar 21.

Carbon dioxide rebreathing in respiratory protective devices: influence of speech and work rate in full-face masks

Carmen L Smith [1], Jane L Whitelaw, Brian Davies

Affiliations

PMID: 23514282 DOI: 10.1080/00140139.2013.777128

Abstract

Carbon dioxide (CO2) rebreathing has been recognised as a concern regarding respirator use and is related to symptoms of discomfort, fatigue, dizziness, headache, muscular weakness and drowsiness. Previous investigations are limited by small sample size and have not evaluated the relationship between CO2 inhalation and phonic respiration (breathing during speech) in respiratory protective devices (RPDs). A total of 40 workers trained in the use of RPDs performed a graded exercise test on a cycle ergonometer that increased in workload every 5 min. During the third minute of each stage, participants read aloud a prepared text. Measures of mixed expired CO2 (PECO2), mixed inspired CO2 (PICO2) and respiration were monitored. The results showed that phonic respiration and low work rates contributed to significantly higher levels of CO2 rebreathing. Aiming to reduce CO2 exposure may result in improved wear time of RPDs. It is recommended that these findings be incorporated in technical specifications regarding human factors for RPDs.

Practitioner summary: Carbon dioxide (CO2) rebreathing in respiratory protective devices (RPDs) has been highlighted as a key concern regarding respirator use. However, the problem is relatively under researched. This paper presents novel findings on the impact of phonic respiration (breathing during speech) and CO2 concentrations in RPDs.

Related information

MedGen
PubChem Compound (MeSH Keyword)

LinkOut - more resources

Full Text Sources
Taylor & Francis

Neurocirugia (Astur). 2008 Apr;19(2):121-6. doi: 10.1016/s1130-1473(08)70235-5.

Preliminary report on surgical mask induced deoxygenation during major surgery

A Beder [1], U Büyükkoçak, H Sabuncuoğlu, Z A Keskil, S Keskil

Affiliations

PMID: 18500410 DOI: 10.1016/s1130-1473(08)70235-5

Abstract

Objectives: This study was undertaken to evaluate whether the surgeons' oxygen saturation of hemoglobin was affected by the surgical mask or not during major operations.

Methods: Repeated measures, longitudinal and prospective observational study was performed on 53 surgeons using a pulse oximeter pre and postoperatively.

Results: Our study revealed a decrease in the oxygen saturation of arterial pulsations (SpO2) and a slight increase in pulse rates compared to preoperative values in all surgeon groups. The decrease was more prominent in the surgeons aged over 35.

Conclusions: Considering our findings, pulse rates of the surgeon's increase and SpO2 decrease after the first hour. This early change in SpO2 may be either due to the facial mask or the operational stress. Since a very small decrease in saturation at this level, reflects a large decrease in PaO2, our findings may have a clinical value for the health workers and the surgeons.

Related information

PubChem Compound (MeSH Keyword)

LinkOut - more resources

Full Text Sources

Elsevier Science

Medical

MedlinePlus Health Information

Ann. Occup. Hyg., Vol. 56, No. 1, pp. 102–112, 2012
Published by Oxford University Press
on behalf of the British Occupational Hygiene Society
doi:10.1093/annhyg/mer069

Review

Protective Facemask Impact on Human Thermoregulation: An Overview

RAYMOND J. ROBERGE*, JUNG-HYUN KIM and AITOR COCA

National Personal Protective Technology Laboratory, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention

Received 6 April 2011; in final form 5 July 2011; published online 13 September 2011

The use of protective facemasks (PFMs) negatively impacts respiratory and dermal mechanisms of human thermoregulation through impairment of convection, evaporation, and radiation processes. The relatively minor reported increases in core temperature directly attributable to the wearing of PFMs suggest that associated perceptions of increased body temperature may have a significant psychological component or that regional or global brain temperature changes are involved. Modifications in PFM structure, components, and materials might allow for improved heat dissipation and enhanced compliance with use.

Keywords: comfort; core temperature; PFMs; thermoregulation; tolerance

INTRODUCTION

The spate of serious viral respiratory infectious agent outbreaks (e.g. severe acute respiratory syndrome, avian influenza, and pandemic influenza) has placed significant impetus upon the use of protective facemasks (PFMs), including filtering facepiece respirators (FFRs), surgical/medical facemasks (FM), and elastomeric air-purifying respirators (EAPRs) by healthcare workers (HCWs) and the public. The most commonly employed PFMs in these situations are FFRs and FMs. FFR are tight-fitting particulate respirators with a filter as an integral part of the facepiece or with the entire facepiece composed of the filtering medium that covers at least the mouth and nose and filters out harmful particles (NIOSH, 2004). FMs are loose-fitting disposable masks that cover the nose and mouth and are referred to by various nomenclatures, such as surgical mask, medical mask, procedure mask, dental mask, and laser mask.

FMs were initially introduced into surgery to not only prevent surgical personnel from contaminating the surgical field with respiratory droplets expelled during speaking, coughing, and sneezing but also protect the wearer from splashes or sprays (TFAH and AAP, 2009). Because of their loose fit, FMs are unable to provide a high degree of protection from airborne particulates of small dimensions (i.e. droplet nuclei) that can harbor pathogens (Oberg and Brosseau, 2008). EAPRs are reusable, air-purifying respirators (APR) with facepieces made of pliable materials (e.g. silicone, rubber, and plastic) that employ one or two particulate cartridge filters and come in full facepiece or half-mask models (Roberge *et al.*, 2010d). Although there is currently some ongoing debate and investigation into the relative merits of FFR versus FM in protecting the wearer from pathogens (Loeb *et al.*, 2009; Srinivasan and Perl, 2009; Gralton and McLaws, 2010), there is less controversy regarding their being of some efficacy in preventing the transmission of respiratory pathogens (Cowling *et al.*, 2009; MacIntyre *et al.*, 2009; Aiello *et al.*, 2010). However, the use of PFMs will not be effective if not used appropriately.

One of the more frequently cited reasons for intolerance and associated lack of compliance with appropriate PFM use is the discomfort related to buildup of facial heat (Jones, 1991; Laird *et al.*, 2002; Radonovich *et al.*, 2009). In a recent study (Baig *et al.*, 2010),

*Author to whom correspondence should be addressed
Tel: +1-412-386-5241; Fax: +1-412-386-6864;
e-mail: dtn0@cdc.gov

increased facial heat was experienced 'frequently-to-always' by 56% of HCWs wearing N95 FFR. PFM-associated facial heat complaints may represent any of a variety of effects, including local dermal effects, increased temperature of breathing air, elevated core temperature, or psychophysiological responses. This review will examine the etiology of PFM-associated increases in the body's heat perception and burden and suggest potential mitigation strategies.

METHODS

A computerized literature search was undertaken for the period 1950–2010 with the search engines Medline®, OvidSP®, EMBase™, PsycINFO®, Compendex®, and Google®. A web-based search of relevant electronic references was also performed and the bibliographies of selected articles and textbooks were perused for pertinent articles (Fig. 1). References selected for inclusion in the review were those that included information relating to heat, comfort, and tolerance associated with the use of PFMs.

RESULTS

A total of 195 articles from the literature was retrieved along with 42 web-based relevant articles and one textbook chapter. Of these, 84 literature references serve as the database for this study, including 80 journal articles, 3 electronic references from medical, governmental, and news agency sources, and 1 book chapter. There is a paucity of data available on the influence of PFMs upon body thermoregulation.

DISCUSSION

The genesis of PFM-associated changes in body temperature is a composite of several inputs of variable prominence that includes respiratory heat exchange mechanisms, the impact of nasal versus oral respiration, metabolic cost and thermal load of PFMs, facial skin heat load of PFMs, ambient climate and PFM microclimate (i.e. PFM dead space) heat and humidity, and psychophysiological heat response components.

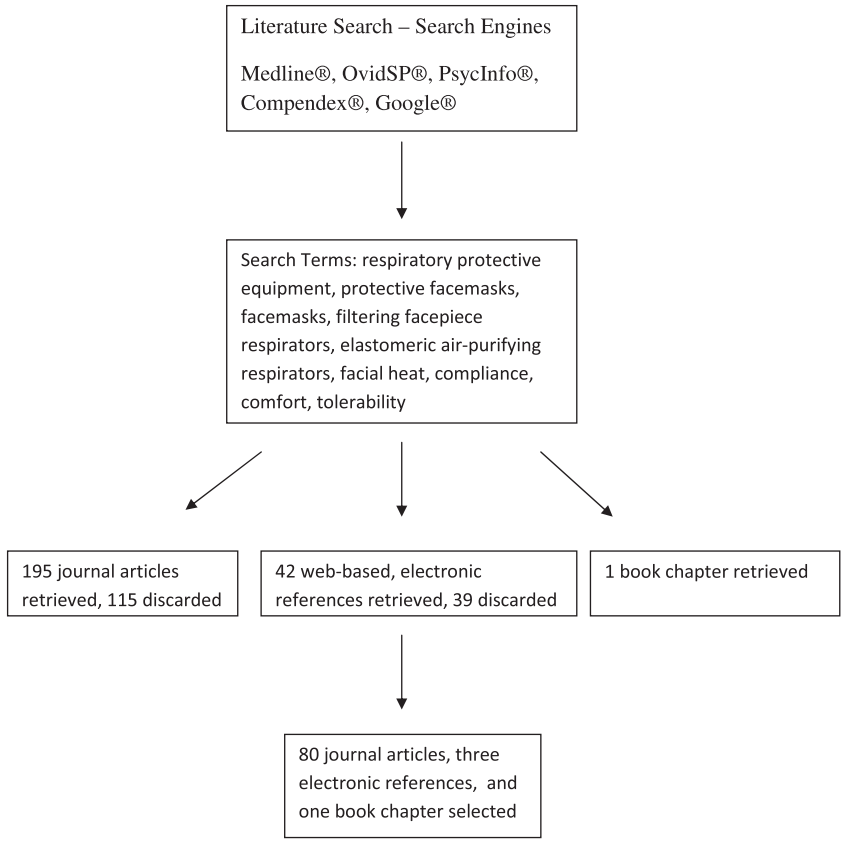


Fig. 1. Literature review data sources.

Respiratory heat exchange mechanisms

Excess heat generated by the body's metabolism and transferred from environmental heat sources (e.g. radiation) must be released to the surrounding environment in order to maintain thermal homeostasis. While human heat balance can be conceptually explained in various forms, the following heat balance equation, re-written from Parsons (2003)' provides a practical approach for its estimation:

$$S = M - W - (C + R + E_{sk}) + (C_{res} + E_{res}),$$

where S = rate of heat storage ($W \cdot m^{-2}$), M = rate of metabolic energy production, W = rate of the body's mechanical work, C = rate of convective heat loss from the skin, R = rate of radiative heat loss from the skin, E_{sk}= rate of evaporative heat loss from the skin, C_{res} = rate of convective heat loss from respiration, and E_{res} = rate of evaporative heat loss from respiration. Thus, the body achieves heat balance when S equals zero. As a point of interest, heat exchange (loss) through respiration consists of two components: convective heat loss as a function of cool air inhalation in which heat from the lungs is transferred in exhalation (C_{res}) and evaporative heat loss as a function of moisture saturation in exhaled air (E_{res}). In practice, the amount of respiratory heat loss can be quantified using the following equation (Parsons, 2003):

$$C_{res} + E_{res} = (0.0014\,M\,[34 - T_a] + 0.0173\,M\,[5.87 - p_a]),$$

where T_a = ambient temperature (°C) and P_a = ambient pressure (kPa). Under thermoneutral environmental conditions, inspired air is warmed and saturated in the lungs which are generally at core temperature, but exhaled air temperature is lower (e.g. ~34–35°C, Tozer, 1924; Winslow *et al.*, 1943) because some warmth and moisture are reclaimed in transit through the nasal passages. The proportion of heat loss through each component of the respiratory heat loss mechanism has not yet been determined, but it is generally agreed that a greater amount of heat is lost through E_{res} than C_{res} due to the fact that the latent heat of water evaporation is much greater than the specific heat of air. While a number of previous investigations have shown that respiratory heat loss is dependent on several variables such as temperature and vapor gradients of inspired air (McCutchan and Taylor, 1951; Cole, 1953), respiratory minute volume (Cole, 1953), changes in body temperature (Hanson, 1974), health status (e.g. asthma) (Burch, 1945; Deal *et al.*, 1979), and working/exercise status (Cain *et al.*, 1990; Livingstone *et al.*, 1994), the total amount of respiratory heat loss as a function of C_{res} and E_{res} under normal condition is 10–15 Watts (W), which accounts for ~10% of total heat loss from the body (Burch, 1945; Ingelstedt, 1956; Hanson, 1974).

Nasal and oral respiratory pathways of thermostasis

The majority of healthy adults are nasal breathers at resting tidal breathing or light exertion (Niinimaa *et al.*, 1980; Hallani *et al.*, 2008), but changes in the partitioning of the breathing cycle among nasal, oronasal, and oral components can impact the respiratory portion of heat exchange, as well as microclimate (i.e. respirator dead space) heat and moisture content. The use of PFMs results in a switch from nasal to oral breathing in most adults (Harber *et al.*, 1997) and respiratory heat exchange is impacted variably by the route of respiration. Nasal breathing is associated with less heat loss to the environment than oronasal and mouth breathing because some expired heat and humidity are reclaimed by the rich vasculature and mucosal surfaces of the nasal passages and paranasal sinuses (Harber *et al.*, 1997; Holden *et al.*, 1999). The nasal mucosa normally recovers one-third of the water delivered to the inspiratory airflow from the expiratory airflow (Martins De Araujo *et al.*, 2000). When the metabolic rate is significant (e.g. during strenuous physical activity), a shift to oronasal breathing occurs that is associated with a greater respiratory minute volume (Niinimaa *et al.*, 1980) and the percentage of mouth breathing increases as the metabolic rate increases (Harber *et al.*, 1997). Increases in core temperature of ~1°C are associated with induction of hyperventilation (increase in pulmonary ventilation of ~35%) relative to metabolic needs (White, 2006). Results from a previous investigation (Varene *et al.*, 1986) showed that the temperature and amount of water delivered on expired air are significantly greater with mouth breathing than nasal breathing. Therefore, there would likely be an increase in the net respiratory heat loss to the environment with oronasal breathing over that noted with nasal breathing only, especially during strenuous physical work and hyperventilation. The net respiratory heat loss through oronasal breathing at a high workload (150 W), under temperate ambient temperature (25°C), has been reported as 103 ± 12 W, which accounts for ~46% of total cephalic heat loss (Rasch *et al.*, 1991).

The contribution of the breathing pathway to central nervous system temperature regulation has been an area of interest for some time. Hirata *et al.* (1978)

observed that tympanic temperatures (considered an indirect measure of core temperature) were consistently higher with mouth breathing, implying that the vascular supply to the head had been cooled more by normal nasal breathing. There have been studies proposing a mechanism of selective brain cooling (SBC) in which venous blood is cooled in the facial area and delivered through a direct venous pathway to the cranium to directly cool the brain and serve as a protective mechanism, especially in hyperthermic states (Cabanac and Caputa, 1979; Cabanac, 1993). This mechanism can be enhanced by nasal breathing and sweat evaporation on the head (Nagasaka *et al.*, 1998). Supporting this concept of SBC is the finding of a cooling effect (0.4–0.8°C) on the frontobasal aspects of the human brain (a site in proximity to the hypothalamus, the major thermoregulatory area of the brain) in post-operative, fully conscious neurosurgical patients with mild hyperthermia spontaneously nasal breathing for 3 min (18–20 breaths min^{-1}) in ambient temperature of 22°C (Mariak *et al.*, 1999). This can perhaps be partially explained by the fact that the distance between the roof of the nose and the floor of the anterior cranial fossa is less than a millimeter (Mariak *et al.*, 1999). Thus, evaporative cooling of the nasal mucosa through intensive nasal breathing directly impacts temperature on the frontobasal aspects of the brain. However, it is worth noting that the issue of whether an effect of SBC is only limited to a local brain region or to the entire brain (which constitutes a significant reduction in thermal gradients of body core temperature) still remains unresolved. Of note, some studies showed that mouth breathing results in a lowering of oral temperature readings due to the cooling effects of ventilatory air on the oral mucosa (Maron, 1983), with significantly lower temperature readings at the anterior sublingual and dorsum of the tongue sites than on the posterior sublingual and buccal trough sites (Cooper and Abrams, 1984). Others have also reported on elevated tympanic temperatures associated with mouth breathing (Neff *et al.*, 1989; Dezell, 1994). Thus, although it is apparent that different anatomic pathways for respiration (i.e. nose, mouth) influence thermoregulation to different degree, the use of multiple types of temperature monitoring methodologies (e.g. oral, tympanic, brain, and skin temperature measurements) reported in the research literature makes it difficult to determine accurately the full impact of nasal and oral respiration on core temperature. In general, based on available data, at low-to-moderate work rates, PFM-related increases in core temperature will likely be minor, irrespective of the route of respiration.

Metabolic cost and thermal load of protective facemasks

The direct contribution of PFMs to the metabolic cost is considered to be minor: PFMs with low/moderate filter performance (typically resulting in lower levels of airflow resistance) [i.e. European classification P1 and P2 filters (80% and 94% filtration, respectively, at test conditions of 95 l min^{-1} constant air flow rate)] add a metabolic cost of 20 W m^{-2}, and for PFMs with high performance filters [i.e. European classification P3 filters (99.95% filtration at test conditions of 95 l min^{-1} constant air flow rate)], the metabolic cost is 40 W m^{-2} (based on 1.8 m^2 body surface area) (Hanson, 1999). This mild effect of PFMs on energy expenditure at low-to-moderate work rates is supported by a recent study of HCWs wearing low-resistance PFMs [i.e. surgical masks and P2 equivalent FFR (i.e. N95 FFR)] during usual work activities for 30 min that showed increases in tympanic temperature of only 0.07 and 0.03°C, respectively (Yip *et al.*, 2005). Similarly, no added metabolic/thermal load was demonstrated for tight-fitting powered air-purifying respirators (PAPR) or a negative pressure full facepiece APR used in warm environments (33.9–35°C dry bulb temperatures) at low–moderate treadmill work rates for 20 min (Caretti, 2002; Caretti and Gardner, 2003). Under high heat/high work conditions (43.3°C/116 W h^{-1}) over 1 h, oral temperature increased only 0.33°C when wearing a full facepiece APR and no significant effect was noted under high heat/low work (58 W h^{-1}), low work/high heat, and low work/low heat (25°C) scenarios, whereas conditions remained basically unchanged with a half facepiece APR (James *et al.*, 1984). At a work rate of 200–300 Kcal h^{-1}, no significant differences were noted in core (rectal) temperature over 2 h for subjects wearing a full facepiece APR compared with not wearing a respirator (Martin and Callaway, 1974). Similarly, Guo *et al.* (2008) also reported that tympanic temperature rose only 0.2°C for FFR with an exhalation valve (N95FFR-EV) and 0.6°C for FFR during staggered treadmill exercise at 3.2 km h^{-1} × 20 min, 4.6 km h^{-1} × 10 min, and 6.4 km h^{-1} × 10 min, with interspersed 10-min rest periods. In another study (Hayashi and Tokura, 2004), tympanic temperatures in four female subjects, at the end of performing 3 series of 15-min stepping exercises interspersed with 5-min rest periods at environmental conditions of 28°C temperature and 60% relative humidity (RH), showed increases ranging from ~0.25 to 0.5°C for N95 FFR-EV and 0.25–1.4°C for N95 FFR; increases in rectal temperature for the same exercise period were ~0.7 and 0.9°C, respectively. However, the timing of the menstrual cycle was not identified which could have

impacted temperature measurements, and the subjects were wearing protective garments (Gore-Tex) which have somewhat limited vapor permeability that could also result in heat retention, so that it is difficult to partition out the FFR component of the rise in temperatures. Thus, it would appear from the limited available data that PFM use for periods $\leq$1 h, under varying workloads (low, moderate, and high), has, in and of itself, limited metabolic impact and is generally associated with only minimal-to-mild increases in body temperature as measured by oral or tympanic routes. While it is a common practice to reference a level of tympanic temperature as core body temperature in determining the thermal impact of PFMs, consideration must be given to the fact that studies have demonstrated that significant variability exists between concurrent measurements at both ears and between tympanic temperatures and pulmonary artery temperature (the 'gold standard' for core temperature) (Fulbrook, 1997; Sanderson *et al.*, 2010). Logically, the impact on body temperature is likely to be augmented with longer, uninterrupted periods of PFM use in high ambient temperatures and humidity and at higher work rates.

Facial skin temperature changes with protective facemasks

The head is an area of very high metabolic activity and is a critical structure for cooling, especially when the remainder of the body is impeded in normal heat dispersal (James *et al.*, 1984). The heat flux per unit area of bare facial skin is 104 W m^{-2}, approximately double the 50 W m^{-2} flux of the rest of the body (DuBois *et al.*, 1990). In moderate environmental conditions, the average temperatures of peripheral tissues are 2–4°C lower than core temperature (Lenhardt and Sessler, 2006). Facial skin temperature in an adult can vary considerably by anatomic region, with the nasolabial and perioral areas (those areas most frequently covered by PFMs) having been reported as having the highest baseline facial temperatures in young adults (34.6 ± 1.7, 34.1°C ± 1.7) and older adults (35.3 ± 1.4, 35.2°C ± 1.3) (Marrakchi and Maibach, 2007). Body temperatures are regulated, in large measure, by the exchange of heat through the body's skin where radiation, convection, and evaporative processes occur, as described earlier. Obviously, in the facial region, these processes can only occur to their optimal extent with adequate facial skin exposure to the ambient environment, a situation that is impeded by the barrier effect of PFMs (Hanson, 1999). PFM facepiece materials and design significantly impact overall comfort (Caretti and Coyne, 2008) and EAPRs, with their larger non-breathable sealing areas, are likely to have a greater impact on facial skin temperature than the more permeable FFRs and FMs. In addition to the barrier effect, the venous flow from the head and face to the cranial cavity that plays a role in brain cooling (Cabanac and Caputa, 1979) could theoretically be compromised by pressure from the straps and head harness of tight-fitting PFs (i.e. EAPRs). It has been posited that the discomfort of respirator wear is related to elevations in facial skin temperature (DuBois *et al.*, 1990). Multiple studies have reported on the impact of PFMs upon facial skin temperature, but most do not report concurrent core temperatures that would assist in clarifying the central versus peripheral impact of PFM use upon body temperature. The contribution of facial skin temperature upon EAPR comfort parameters at 25°C ambient temperature is found in a formula derived by linear regression analysis of data from multiple studies (Caretti and Coyne, 2008):

$$\begin{aligned}\text{Comfort} \leq 25 = 0.59 &+ \left(0.06 \times \text{TS}_{\text{face}}\right) \\ &+ (0.20 \times \text{facepiece}) + (0.29 \times \text{nose cup}) \\ &+ (0.25 \times \text{harness}) + (0.22 \times \text{breathing}),\end{aligned}$$

where TS_{face} = thermal sensation of the face, facepiece is a subjective rating of facepiece comfort (unit less), nose cup is subjective rating of nose cup comfort (unit less), harness is head harness comfort rating (unit less), and breathing is breathing comfort score (unit less).

At >25°C, the formula is: Comfort >25 = 0.40 + (0.12 ×TS_{face}) + (0.17 × facepiece) + (0.32 × nose cup) + (0.17 × harness) + (0.36 × breathing).

At ambient temperatures of 18.9–25.5°C and 49–63% RH, skin temperatures at the tip of the nose and at the chin increased 3.7–7.3 and 2.6–3.6°C, respectively, during sedentary activity while wearing surgical masks over a 15-min period (Enerson *et al.*, 1967). Laird *et al.* (2002) reported that wearing a filter-type respirator during the last 15 min of a 30-min laboratory study at a low work rate (50 W) resulted in a 1.9°C increase in upper lip temperature, but had no effect on cheek temperatures not covered by the respirator. The second portion of that study, a simulated work environment at ambient temperatures of 17–24°C and 60–80% RH, led to increases in upper lip temperature of 0.5–2.4°C, but again had no effect on cheek temperatures measured outside the respirator. Johnson *et al.* (1997) noted that, at ambient conditions of 35°C and 90% RH, and sedentary activity for 90 min, skin temperature under a full facepiece APR rose by 2°C. Over a 30-min period at ambient temperatures of 21–26°C, skin temperatures taken

under rubber EAPRs and dust-mist fiber masks (a type of FFR in common use before 1995) rose 1.5 and 1.1°C, respectively, above baseline values (DuBois *et al.*, 1990). From a subjective perception, PFM-related upper lip temperatures >34°C elicited sensations of warmth (and associated discomfort), whereas those below this level were sensed as cool-to-neutral (Gwosdow *et al.*, 1989; DuBois *et al.*, 1990), despite the fact that these temperatures are within the realm of normal facial temperatures. However, consideration must be given to the 90–100% RH levels attained in PFM that result in a PFM microenvironment heat index (combination of temperature and humidity effects) that may be quite high. For example, at air temperature of 34°C and 95% RH (e.g. equivalent to expired air under normal conditions), the PFM microenvironment heat index could be 62°C on exhalation, though it would subsequently be diminished variably by the admixture of inhaled ambient air.

Studies documenting the effect of PFMs upon facial skin and core temperatures concurrently are rare. Following 3 series of 15-min stepping exercises interspersed with 5-min rest periods at ambient temperature of 28°C and 60% RH, cheek temperatures under N95 FFR and N95 FFR-EV rose ∼2.0 and 1.5°C, respectively, with concomitant increases in rectal temperature ranging from ∼0.6 to 1.1°C and tympanic temperature increases from 0.3 to 1.3°C, but the subjects were wearing protective clothing ensembles (Gore-Tex) that could have added to the heat load (Hayashi and Tokura, 2004). It is perhaps not surprising that increases in temperature of skin under PFMs would not have a dramatic impact on core temperature given that PFMs cover a portion of the face that accounts for only 1–2% of body surface area, so that the amount of heat transfer to the core from this heated facial skin should only approximate similar percentages (McCaffrey *et al.*, 1975). Importantly, tympanic temperature measurements cannot be relied upon as accurate indicators of central blood temperature because they are susceptible to modification by the local environment such as when localized regions of heating are present on the face (e.g. when wearing PFMs) (McCaffrey *et al.*, 1975) or when the face is cooled (Shiraki *et al.*, 1988). Thus, all forms of negative pressure PFMs elevate the underlying skin temperature to variable degree based upon the PFM type, fit (gaps in the seal might allow for more cooling), composite materials (e.g. silicone, polypropylene fibers, etc.), work rate, ambient conditions, and duration of use. However, this effect is noted only for skin covered by the PFM and does not seemingly impact the facial skin that is not covered; uncovered facial skin mean temperature is a linear function of ambient temperature (Nielsen *et al.*, 1987a). Conversely, PAPR have been shown to actually decrease core temperature due to the cooling effects of their fan-supplied air (Caretti and Gardner, 2003). The limited currently available data do not allow for determination of any distinct correlation between elevated facial skin temperature underneath PFMs and concurrent core temperature, but the small area of the face covered by PFMs suggests that its contribution to core temperature would not be excessive, but may have a significant impact on the perception of thermal comfort.

PFM dead space heat and humidity

Facial skin temperatures are impacted by the temperature and humidity of the surrounding air under normal conditions (Nielsen *et al.*, 1987b). When ambient temperatures are lower than facial skin temperature, radiation is the main source of heat loss. In hot conditions, especially combined with significant physical activity, when temperatures approach or exceed body temperature, evaporative cooling (sweat evaporation) becomes a dominant heat exchange mechanism (Hanson, 1999). Wearing PFMs creates a microenvironment (i.e. PFM dead space) that then becomes the wearer's breathing environment. This microenvironment has a significant impact on heat exchange processes of the facial skin. PFM microenvironment temperature has been considered a key parameter indicating thermal stress (Li *et al.*, 2005). In ambient conditions of high temperatures, the dissipation of heat from the PFM dead space can be negatively impacted due to a decreased temperature gradient between the ambient environment and the PFM microenvironment (Li *et al.*, 2005). The PFM dead space 'effective temperature' (a single quantitative index of environmental discomfort that incorporates air temperature and humidity) can be quite high. The relatively high heat and humidity of the expired air can cause moisture to condense on the outer surface of the FFR due to the temperature difference between the FFR and the environment (Li *et al.*, 2006). This phenomenon can negatively impact the vapor and air permeability of the FFR, which consequently impairs respiratory heat loss and imposes an increased heat burden. Consideration must also be given to the amount of sweat formed within the dead space of PFMs. Sweat rates for the head, face, and neck averaged 0.203 gm^{-1} min^{-1} sedentary while wearing a full facepiece APR in a warm humid environment (35°C, 90% RH) over 90 min, but most of the sweat came from the neck. It was estimated that 7.5 gm h^{-1} of sweat could accumulate in the respirator (Johnson *et al.*, 1997). At wet bulb globe

temperature of 19.3°C and moderate treadmill exercise at 75% maximum heart rate while wearing a full facepiece EAPR, facial sweat was 1.05 g min^{-1} (Caretti and Gardner, 1999). Increased retention of water vapor and sweat within PFMs has other important ramifications in addition to effects on comfort because it can affect the facial seal of the PFMs (Caretti and Gardner, 1999), potentially increase the breathing resistance (Roberge *et al.*, 2010a), and theoretically increase the risk of transmission of infectious agents to the wearer via a wicking mechanism (Yi *et al.*, 2005). However, recent studies have demonstrated that, over the course of 1 h of low work rate exercise, FFR with and without an exhalation valve and EAPR with an exhalation valve retained very little moisture, which was attributed to the use of hydrophobic fibers (i.e. polypropylene) and exhalation valves, as well as the use of low work rates in some studies (Roberge *et al.*, 2010b,c,d). Thus, at low-to-moderate work rates, the microenvironment of PFMs develops mild-to-moderate increases in temperature with concurrent high humidity levels that increase the effective temperature to uncomfortable levels, impact comfort and tolerance, and potentially reduce respiratory heat exchange.

Psychophysiological heat responses

The face is relatively uniform in its sensitivity to warming when compared to the mouth (Green and Gelhard, 1987), but the area of the face that is covered by PFMs is very thermosensitive (Laird *et al.*, 1999). This may be possibly due to a higher facial thermoreceptor density, as has been demonstrated in animals (Cheung, 2010). The microenvironment air temperature increases the temperature of facial skin covered by PFMs that, in turn, significantly influences thermal sensations of the whole body, a phenomenon that may have a neurological component that has been explained as being due to the possibility that afferent impulses from the face to the central nervous system may be weighted more than those from other areas (Nielsen *et al.*, 1987b). Also, the impairment of heat exchange in the facial and head regions may have a more profound impact given the fact that these areas are so critical for thermal regulation (James *et al.*, 1984). The highly thermosensitive nature of the face is exemplified by the fact that cooling of the face is two to five times more effective at suppressing sweating and thermal discomfort than cooling an equivalent dermal area elsewhere on the body (Cotter and Taylor, 2005).

Purely psychological phenomena can indirectly impact the thermal load associated with PFM use. Individuals with underlying anxiety disorders (e.g. panic attacks) are at risk of provoking same when wearing PFMs. The respiratory subtype of panic disorder displays prominent respiratory symptomatology during panic attacks that is probably linked to a false suffocation alarm in the central nervous system (Freire *et al.*, 2010). Individuals with panic disorder are deemed to be very sensitive to increases in CO_2 levels in the body and PFM use is associated with retention of CO_2 in some individuals (Roberge *et al.*, 2010c,d) that could potentially serve as a trigger to a panic attack (Morgan, 1983). Indeed, single breath 35% CO_2 inhalation is a standard provocation test for panic disorder (Valenca *et al.*, 2002). Wearing PFM (e.g. gas masks) can cause claustrophobic sensations and has been used as a provocative maneuver in mild-to-moderate cases of claustrophobia (Rachman, 1993; Radomsky *et al.*, 2001). The usual response to the onset of a panic attack or claustrophobic reaction, irrespective of the triggering event, is a sympathomimetic one brought about by the release of neurotransmitters (e.g. catecholamines such as adrenalin and noradrenalin). Release of these neurotransmitters results in increased metabolic activity manifested physically as elevated heart rate and respiratory rate, palpitations, elevated blood pressure, etc., the so-called 'fight or flight' phenomenon. An associated sensation of warmth in these events may be due to actual increases in body temperature brought about by the increase in metabolic activity, by neurosensory phenomena (flushing of the skin), increased respiratory effort associated with overcoming perceived increases in PFM-related breathing resistance, or by increased sweating in the PFM microenvironment brought about by psychological stress that could increase the effective temperature of that area of the face. It may be that, in temperate environments, some (possibly significant) portion of the sensation of excess heat and warmth associated with the use of PFMs has a psychological basis given that the PFM metabolic and facial heat contributions themselves are not excessive. Much of the available research supports the notion that the primary thermal effect of wearing a respirator is subjective discomfort (Caretti and Coyne, 2008). Conversely, increased body temperature associated with thermal stress can itself lead to decrements in psychomotor performance in those with no recognized psychopathology (Morgan, 1983). The psychology of PFM use has received some limited investigation in the past and would benefit from significantly more study.

Potential mitigation strategies for protective facemask-associated heat retention

Mitigation of PFM-associated heat is desirable for comfort that results in greater PFM tolerance and ultimately translates to greater protection for the wearer.

Some strategies aimed at decreasing the heat burden attributed to PFMs could be explored, including (but not limited to):

(i) 'Promotion of nasal breathing when wearing PFMs'—Because nasal breathing likely results in less heat and humidity retained in the microclimate of PFMs and may have favorable effects on cooling of some brain structures, it may be desirable from a thermal perspective to promote nasal breathing through the education of wearers of PFMs. This would be feasible only for low-to-moderate work rates, as higher energy expenditures cause a switch to oronasal breathing (Harber *et al.*, 1997).

(ii) 'Investigation of the effect(s) of pre-use refrigeration of PFMs on facial and body temperatures—It has been anecdotally mentioned that cooling of EAPR might be a simple method of decreasing the impact of heat on wear (Laird *et al.*, 2002). Although silicon and rubber used to construct the body of EAPR could be cooled in such a manner, research exists neither on the length of duration of the cooling effects nor on the impact of cooling on the fit of the PFMs. Future research could be directed toward identifying PFM-compatible materials with cooling-retention features, especially in light of the fact that facial cooling is two to five times as effective at reducing thermal discomfort than equivalent areas of the skin in other body regions (Cotter and Taylor, 2005).

(iii) 'Use of exhalation valves'—PFMs with exhalation valves are touted as increasing wearer comfort through facilitated dispersal of PFM dead space heat and humidity to the environment. However, at the low and moderate work rates that most current workers experience (Meyer *et al.*, 1997; Harber *et al.*, 2009), the benefits of exhalation valves (in FFR) may not be realized because the development of the requisite streamlined air currents to activate the valve may not occur (Roberge *et al.*, 2010c,d) as it does with EAPR. Improvements in design and function could potentially lead to exhalation valves that function with lesser airflow gradients that might afford greater heat and humidity losses at lower energy expenditures.

(iv) 'Investigate the breathability of PFM filtration materials'—While there is likely a tradeoff between breathability (vapor and air permeability) and PFM filtration efficiency (that is critical to reducing the risk of exposure to harmful particles and infectious agents), it would be of importance to fully investigate the material properties of PFMs to ensure optimal breathability that could lead to subsequent lowering of PFM dead space humidity levels that impact comfort and tolerance. For example, nanofibers offer filtration efficiency with a concomitant decrease in breathing resistance over that noted with other meltblown and spunbonded filter materials (Qion and Wang, 2006; Lee and Obendorf, 2007).

(v) 'Development of PFMs with miniaturized battery-powered fans'—Fan-derived air currents, as exemplified by PAPRs and surgical hoods, cool the head and facial regions and the inhaled air resulting in minimal increase or a decrease in body temperature (Caretti and Gardner, 2003). Miniaturized ($8 \times 8 \times 3$ mm) battery powered fans currently exist for cooling various small electronic appliances (e.g. smart phones, GPs modules, etc.) and could potentially be adapted to PFMs (http://www.sunonamerica.com/pdf/mm_fan_catalog.pdf). One such model currently on the market, the BL-50 from Koken, Ltd. (Tokyo, Japan), is a half-mask that contains a battery-powered integral blower triggered by inhalation and used to maintain constant pressure within the facepiece (Richardson and Hofacre, 2008). In addition to cooling the face, the development of positive pressure by integral fans could serve to enhance respiratory protection by preventing ingress of harmful particles or organisms into the PFM.

(vi) 'PFM dead space parameters'—Re-breathing of retained warm expired air within the dead space of PFMs increases facial heat discomfort. Some styles of PFM (e.g. cup shaped and duckbill FFR and FM) have larger dead spaces and can thus result in greater volumes of retained warmed air than other styles (e.g. flat fold and pleated FFR and FM). A recent study of a cohort of HCWs using PFM reported that 81% of HCWs interviewed used either a cup-shaped or duckbill N95 FFR and that 56% of all interviewees indicated that they experienced increased facial heat 'most or all of the time' (Baig *et al.*, 2010). Therefore, it would be important to study the impact of various styles of PFMs on facial heat in order to determine those styles with lower associated increases in facial heat.

(vii) Anxiety-related perceptions of PFM-associated heat—The retention of CO_2 with the use of PFMs is a possibility (Roberge *et al.*, 2010d), and panic disorder can be triggered by elevated CO_2 levels. Some of the symptoms of panic disorder include hot flashes and sweating.

Response to the 35% single breath CO_2 inhalation provocation test is quite specific for panic disorder and is routinely utilized for this diagnosis. Individuals who demonstrate intolerance to PFMs could undergo non-invasive transcutaneous CO_2 monitoring and a CO_2 provocation test to assist in determining if CO_2 retention is the source of their symptoms.

The large number of PFM users (private industry, HCWs, the public) and the increased use of PFMs in certain scenarios (e.g. infectious agent outbreaks, environmental disaster remediation efforts, etc.) should make PFM-related effects on thermoregulation a major focus for researchers and should serve as a significant impetus for additional investigation. Intolerance to the thermal effects of PFMs leads to decreased use and concomitant decreased protection for the user.

FUNDING

National Personal Protective Technology Laboratory internal operating funds.

Acknowledgements—The authors thank Jay Parker, MS, William Newcomb, Dr W. Jon Williams, and Dr Ronald Shaffer for their manuscript review and insightful comments.

Disclaimer—The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

REFERENCES

Aiello AE, Murray GF, Perez V *et al*. (2010) Mask use, hand hygiene, and seasonal influenza-like illness among young adults: a randomized intervention trial. J Infect Dis; 201: 491–8.

Baig AS, Knapp C, Eagan AE *et al*. (2010) Health care workers' views about respirator use and features that should be included in the next generation of respirators. Am J Infect Control; 38: 18–25.

Burch GE. (1945) The rate of water loss from the respiratory tract of man living in a subtropical climate. Science; 102: 619–20.

Cabanac M. (1993) Selective brain cooling in humans: ''fancy'' or fact? FASEB J; 7: 1143–6;discussion 1146–7.

Cabanac M, Caputa M. (1979) Natural selective cooling of the human brain: evidence of its occurrence and magnitude. J Physiol; 286: 255–64.

Cain JB, Livingstone SD, Nolan RW *et al*. (1990) Respiratory heat loss during work at various ambient temperatures. Respir Physiol; 79: 145–50.

Caretti DM. (2002) Assessment of the thermal load attributable to protective masks. Aberdeen, MD: US Army Edgewood Chemical Biological Center, Aberdeen Proving Ground Report AMSSB-RRT-PR (E5604).

Caretti DM, Coyne KM. (2008) Reassessment of human performance parameter estimates for respiratory protection design and development. Aberdeen, MD: US Army Edgewood Chemical Biological Center Report ECBC-TR-605.

Caretti DM, Gardner PD. (1999) Respirator fit factor performance while sweating. Am Ind Hyg Assoc J; 60: 84–8.

Caretti DM, Gardner PD. (2003) Quantifying the heat stress attributable to respirator wear. J Int Soc Resp Protect; 20: 110–24.

Cheung SS. (2010) Interconnections between thermal perception and exercise capacity in the heat. Scand J Med Sci Sports; 20 (Suppl. 3): 53–9.

Cole P. (1953) Some aspects of temperature, moisture and heat relationship in the upper respiratory tract. J Laryngol Otol; 67: 449–56.

Cooper KH, Abrams RM. (1984) Attributes of the oral cavity as a site for basal body temperature measurements. JOGN Nurs; 13: 125–9.

Cotter JD, Taylor NA. (2005) The distribution of cutaneous sudomotor and alliesthesial thermosensitivity in mildly heat-stressed humans: an open-loop approach. J Physiol; 565: 335–45.

Cowling BJ, Chan KH, Fang VJ *et al*. (2009) Facemasks and hand hygiene to prevent influenza transmission in households: a cluster randomized trial. Ann Intern Med; 151: 437–46.

Deal EC, Jr., McFadden ER, Jr., Ingram RH, Jr. *et al*. (1979) Role of respiratory heat exchange in production of exercise-induced asthma. J Appl Physiol; 46: 467–75.

Dezell N. (1994) Comparison of tympanic, oral and axillary temperatures in women during labor. Department of the Air Force, Wright-Patterson Air Force Base Report 94–22802. Available at http://www.dtic.mil/cgibin/GetTRDoc?AD=ADA281978&Location=U2&doc=GetTRDoc.pdf. Accessed 1 October 2010.

DuBois AB, Harb ZF, Fox SH. (1990) Thermal discomfort of respiratory protective devices. Am Ind Hyg Assoc J; 51: 550–4.

Enerson DM, Eisenfeld LI, Kajikuri H. (1967) Heat and moisture trapping beneath surgical face masks: a consideration of factors affecting the surgeon's discomfort and performance. Surgery; 62: 1007–16.

Freire R, Perna G, Nardi A. (2010) Panic disorder respiratory subtype: psychopathology, laboratory challenge tests, and response to treatment. Harv Rev Psychiatry; 18: 220–9.

Fulbrook P. (1997) Core body temperature measurement: a comparison of axilla, tympanic membrane and pulmonary artery blood temperature. Intensive Crit Care Nurs; 13: 266–72.

Gralton J, McLaws ML. (2010) Protecting healthcare workers from pandemic influenza: N95 or surgical masks? Crit Care Med; 38: 657–67.

Green BG, Gelhard B. (1987) Perception of temperature on oral and facial skin. Somatosens Res; 4: 191–200.

Guo YP, Yi L, Tokura H *et al*. (2008) Evaluation on masks with exhaust valves and with exhaust holes from physiological and subjective responses. J Physiol Anthropol; 27: 93–102.

Gwosdow AR, Nielsen R, Berglund LG *et al*. (1989) Effect of thermal conditions on the acceptability of respiratory protective devices on humans at rest. Am Ind Hyg Assoc J; 50: 188–95.

Hallani M, Wheatley JR, Amis TC. (2008) Initiating oral breathing in response to nasal loading: asthmatics versus healthy subjects. Eur Respir J; 31: 800–6.

Hanson MA. (1999) Development of a draft British standard: the assessment of heat strain for workers wearing personal protective equipment. Ann Occup Hyg; 43: 309–19.

Hanson RDG. (1974) Respiratory heat loss at increased core temperature. J Appl Physiol; 37: 103–7.

Harber P, Beck J, Luo J. (1997) Study of respirator effect on nasal-oral flow partition. Am J Ind Med; 32: 408–12.

Harber P, Bansal S, Santiago S *et al*. (2009) Multidomain subjective response to respirator use during simulated work. J Occup Environ Med; 51: 38–45.

Hayashi C, Tokura H. (2004) The effects of two kinds of mask (with or without exhaust valve) on clothing microclimates inside the mask in participants wearing protective clothing for spraying pesticides. Int Arch Occup Environ Health; 77: 73–8.

Hirata K, Nagasaka T, Sugano Y. (1978) Effect of alternating respiratory pathway on respiratory capacity, and tympanic and forehead skin temperature during exercise. J Aerospace Environ Med; 15: 8–13.

Holden WE, Wilkins JP, Harris M *et al*. (1999) Temperature conditioning of nasal air: effects of vasoactive agents and involvement of nitric oxide. J Appl Physiol; 87: 1260–5.

Ingelstedt S. (1956) Studies on the conditioning of air in the respiratory tract. Acta Otolaryngol Suppl; 131: 1–80.

James R, Dukes-Dobos F, Smith R. (1984) Effects of respirators under heat/work conditions. Am Ind Hyg Assoc J; 45: 399–404.

Johnson AT, Scott WH, Coyne KM *et al*. (1997) Sweat rate inside a full-facepiece respirator. Am Ind Hyg Assoc J; 58: 881–4.

Jones JG. (1991) The physiological cost of wearing a disposable respirator. Am Ind Hyg Assoc J; 52: 219–25.

Laird IS, Pack R, Carr D. (1999) A survey on the use and non-use of respiratory protective equipment in workplaces in a provincial New Zealand city. Ann Occup Hyg; 37: 367–76.

Laird IS, Goldsmith R, Pack RJ *et al*. (2002) The effect on heart rate and facial skin temperature of wearing respiratory protection at work. Ann Occup Hyg; 46: 143–8.

Lee S, Obendorf SK. (2007) Transport properties of layered fabric systems based on electrospun nanofibers. Fibers Polymers; 8: 501–6.

Lenhardt R, Sessler DI. (2006) Estimation of mean body temperature from mean skin and core temperature. Anesthesiology; 105: 1117–21.

Li Y, Tokura H, Guo YP *et al*. (2005) Effects of wearing N95 and surgical facemasks on heart rate, thermal stress and subjective sensations. Int Arch Occup Environ Health; 78: 501–9.

Li Y, Wong T, Chung J *et al*. (2006) In vivo protective performance of N95 respirator and surgical facemask. Am J Ind Med; 49: 1056–65.

Livingstone SD, Nolan RW, Cain JB *et al*. (1994) Effect of working in hot environments on respiratory air temperatures. Eur J Appl Physiol Occup Physiol; 69: 98–101.

Loeb M, Dafoe N, Mahony J *et al*. (2009) Surgical mask vs N95 respirator for preventing influenza among health care workers: a randomized trial. JAMA; 302: 1865–71.

MacIntyre CR, Cauchemez S, Dwyer DE *et al*. (2009) Face mask use and control of respiratory virus transmission in households. Emerg Infect Dis; 15: 233–41.

Mariak Z, White MD, Lewko J *et al*. (1999) Direct cooling of the human brain by heat loss from the upper respiratory tract. J Appl Physiol; 87: 1609–13.

Maron M. (1983) Effect of tachypnea on estimation of body temperature [correspondence]. N Engl J Med; 309: 612–13.

Marrakchi S, Maibach HI. (2007) Biophysical parameters of skin: map of human face, regional, and age-related differences. Contact Dermatitis; 57: 28–34.

Martin H deV, Callaway S. (1974) An evaluation of the heat stress of a protective face mask. Ergonomics; 17: 221–31.

Martins De Araujo MT, Vieira SB, Vasquez EC *et al*. (2000) Heated humidification or face mask to prevent upper airway dryness during continuous positive airway pressure therapy. Chest; 117: 142–7.

McCaffrey TV, McCook RD, Wurster RD. (1975) Effect of head skin temperature on tympanic and oral temperature in man. J Appl Physiol; 39: 114–18.

McCutchan JW, Taylor CL. (1951) Respiratory heart exchange with varying temperature and humidity of inspired air. J Appl Physiol; 4: 121–35.

Meyer JP, Hery M, Herrault J *et al*. (1997) Field study of subjective assessment of negative pressure half-masks. Influence of the work conditions on comfort and efficiency. Appl Ergon; 28: 331–8.

Morgan WP. (1983) Psychological problems associated with the wearing of industrial respirators: a review. Am Ind Hyg Assoc J; 44: 671–6.

Nagasaka T, Brinnel H, Hales JR *et al*. (1998) Selective brain cooling in hyperthermia: the mechanisms and medical implications. Med Hypotheses; 50: 203–11.

Neff J, Ayoub J, Longman A *et al*. (1989) Effect of respiratory rate, respiratory depth, and open versus closed mouth breathing on sublingual temperature. Res Nurs Health; 12: 195–202.

Nielsen R, Berglund LG, Gwosdow AR *et al*. (1987a) Thermal sensation of the body as influenced by the thermal microclimate in a face mask. Ergonomics; 30: 1689–703.

Nielsen R, Gwosdow AR, Berglund LG *et al*. (1987b) The effect of temperature and humidity levels in a protective mask on user acceptability during exercise. Am Ind Hyg Assoc J; 48: 639–45.

Niinimaa V, Cole P, Mintz S *et al*. (1980) The switching point from nasal to oronasal breathing. Respir Physiol; 42: 61–71.

NIOSH. (2004) National Institute for Occupational Safety and Health (NIOSH) Respirator Selection Logic. Available at http://www.cdc.gov/niosh/docs/2005-100/pdfs/05-100.pdf. Accessed 3 September 2010.

Oberg T, Brosseau LM. (2008) Surgical mask fit and filter performance. Am J Infect Control; 36: 276–82.

Parsons K. (2003) Human thermal environments: the effects of hot, moderate, and cold environments on human health, comfort, and performance. 2nd edn. New York: Taylor & Francis. pp. 17.

Qion X-H, Wang S-Y. (2006) Filtration properties of electrospinning nanofibers. J Appl Polymer Sci; 102: 1285–90.

Rachman S. (1993) Analyses of claustrophobia. J Anx Disorders; 7: 281–91.

Radomsky AS, Rachman S, Thordarson DS *et al*. (2001) The Claustrophobia questionnaire. J Anxiety Disord; 15: 287–97.

Radonovich LJ, Jr., Cheng J, Shenal BV *et al*. (2009) Respirator tolerance in health care workers. JAMA; 301: 36–8.

Rasch W, Samson P, Cote J *et al*. (1991) Heat loss from the human head during exercise. J Appl Physiol; 71: 590–5.

Richardson AW, Hofacre KC. (2008) Technology survey for enhancement of chemical, biological, radiological, and nuclear respiratory protection. Aberdeen, MD: Edgewood Chemical Biological Center. Report ECBC-CR-094.

Roberge RJ, Bayer E, Powell JB *et al*. (2010a) Effect of exhaled moisture on breathing resistance of N95 filtering facepiece respirators. Ann Occup Hyg; 54: 671–7.

Roberge RJ, Coca A, Williams WJ *et al*. (2010b) Surgical mask placement over N95 filtering facepiece respirators: physiological effects on healthcare workers. Respirology; 15: 516–21.

Roberge RJ, Coca A, Williams WJ *et al*. (2010c) Physiological impact of the N95 filtering facepiece respirator on healthcare workers. Respir Care; 55: 569–77.

Roberge RJ, Coca A, Williams WJ *et al*. (2010d) Reusable elastomeric air-purifying respirators: physiologic impact on health care workers. Am J Infect Control; 38: 381–6.

Sanderson B, Lim L, Lei K *et al*. (2010) A comparison of core and tympanic temperature measurement in the critically ill. Crit Care; 12 (Suppl. 1): 329.

Shiraki K, Sagawa S, Tajima F *et al*. (1988) Independence of brain and tympanic temperatures in an unanesthetized human. J Appl Physiol; 65: 482–6.

Srinivasan A, Perl TM. (2009) Respiratory protection against influenza. JAMA; 302: 1903–4.

TFAH and AAp. (2009) Pandemic influenza: warning, children at-risk. Available at http://healthyamericans.org/reports/fluchildren/KidsPandemicFlu.pdf. Accessed 21 August 2010.

Tozer P. (1924) The temperature and saturation of exhaled air in relation to catarrhal infections. Br Med J; 1: 1127–9.

Valenca AM, Nardi AE, Nascimento I *et al*. (2002) Respiratory panic disorder subtype and sensitivity to the carbon dioxide challenge test. Braz J Med Biol Res; 35: 783–8.

Varene P, Ferrus L, Manier G *et al*. (1986) Heat and water respiratory exchanges: comparison between mouth and nose breathing in humans. Clin Physiol; 6: 405–14.

White MD. (2006) Components and mechanisms of thermal hyperpnea. J Appl Physiol; 101: 655–63.

Winslow C, Herrington L, Nelbach J. (1943) The influence of atmospheric temperature and humidity upon the dryness of the oral mucosa. Am J Hyg; 35: 27–39.

Yi L, Fengshi L, Qingyon Z. (2005) Numerical simulation of virus diffusion in facemask during breathing cycles. Int J Heat Mass Transf; 48: 4229–42.

Yip W, Leung L, Lau P *et al*. (2005) The effect of wearing a face mask on body temperature. Hong Kong J Emer Med; 12: 23–7.

Behavioral/Cognitive

Amygdala Responsivity to High-Level Social Information from Unseen Faces

Jonathan B. Freeman,[1,2] **Ryan M. Stolier,**[1,2] **Zachary A. Ingbretsen,**[2,3] **and Eric A. Hehman**[1,2]

[1]Department of Psychology, New York University, New York, New York 10003, [2]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, New Hampshire 03755, and [3]Department of Psychology, Harvard University, Cambridge, Massachusetts 02138

Previous research shows that the amygdala automatically responds to a face's trustworthiness when a face is clearly visible. However, it is unclear whether the amygdala could evaluate such high-level facial information without a face being consciously perceived. Using a backward masking paradigm, we demonstrate in two functional neuroimaging experiments that the human amygdala is sensitive to subliminal variation in facial trustworthiness. Regions in the amygdala tracked how untrustworthy a face appeared (i.e., negative-linear responses) as well as the overall strength of a face's trustworthiness signal (i.e., nonlinear responses), despite faces not being subjectively seen. This tracking was robust across blocked and event-related designs and both real and computer-generated faces. The findings demonstrate that the amygdala can be influenced by even high-level facial information before that information is consciously perceived, suggesting that the amygdala's processing of social cues in the absence of awareness may be more extensive than previously described.

Key words: **amygdala; backward masking; face; fMRI; social cues**

Introduction

With only a glance, humans instantly form impressions of another's face. Such impressions occur spontaneously and are often beyond our conscious control (Zebrowitz and Montepare, 2008). They help us distinguish friend from foe, or those whom we should trust from those of whom we should be wary. Indeed, a mere 50 ms exposure to a face permits trait inferences that are highly correlated among multiple perceivers, indicating that facial cues provide reliable signals about another's underlying disposition (Bar et al., 2006; Willis and Todorov, 2006).

Previous behavioral studies suggest that face-based evaluations are underpinned by two fundamental dimensions, trustworthiness and dominance (Oosterhof and Todorov, 2008). Facial trustworthiness in particular accounts for the bulk of variance in social evaluation (Oosterhof and Todorov, 2008), and recent behavioral studies have provided preliminary evidence that individuals might be sensitive to trustworthiness without perceptual awareness (Todorov et al., 2009; Stewart et al., 2012). However, the neural basis of evaluating high-level social information such as trustworthiness from a face outside awareness remains unexplored.

Evaluations of trustworthiness reflect more general face valence and show correlated activity in the amygdala (Winston et al., 2002; Engell et al., 2007), a subcortical region involved in processing the affective significance of social stimuli and important for a variety of social and emotional behaviors (Phelps and LeDoux, 2005; Adolphs, 2010). Consistent with the sensitivity of the amygdala to negative, threat-related stimuli, initial studies reported linear effects, with amygdala activation increasing for faces appearing less trustworthy (Winston et al., 2002; Engell et al., 2007). This effect held true regardless of task demands (Engell et al., 2007; Todorov and Engell, 2008; Todorov et al., 2011), suggesting that the amygdala may code trustworthiness implicitly when a face is clearly visible. More recent studies have reported quadratic effects, with amygdala activation increasing for faces appearing either more or less trustworthy, relative to neutral (Said et al., 2010; Todorov et al., 2011), potentially reflecting the coding of the salience or motivational relevance of a stimulus derived from a face's trustworthiness (Todorov et al., 2013). A recent meta-analysis found that both linear and nonlinear responses to trustworthiness coexist in different amygdala subregions (Mende-Siedlecki et al., 2013). However, it is currently unknown whether the amygdala is sensitive to trustworthiness before a face can reach perceptual awareness.

Just as the detection of threat-relevant stimuli by the amygdala is functionally adaptive (Öhman, 2005), so too might be its evaluation of a face's trustworthiness. Indeed, a large percentage of death throughout human history is a result of tribal conflicts and coalitional aggression, estimated to have had a substantial impact on human evolution (Chagnon, 1988). Thus, automatic evaluation of another's likelihood to harm or help via facial trustworthiness would facilitate survival and resource maintenance (McDonald et al., 2012). By facilitating the coding of another's trustworthiness before awareness, the amygdala could modulate cortical processes and motivate appropriate behavioral responses. Here, we describe two functional neuroimaging experiments designed to test whether the amygdala can respond to a

Received Dec. 4, 2013; revised June 26, 2014; accepted June 27, 2014.

Author contributions: J.B.F. designed research; R.M.S., Z.A.I., and E.A.H. performed research; J.B.F. and R.M.S. analyzed data; J.B.F. wrote the paper.

We thank Paul Whalen and Todd Heatherton for helpful comments.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Jonathan B. Freeman, Department of Psychology, New York University, 6 Washington Place, New York, NY 10003. E-mail: jon.freeman@nyu.edu.

DOI:10.1523/JNEUROSCI.5063-13.2014

 0270-6474/14/3410573-09$15.00/0

face's trustworthiness without perceptual awareness and to characterize the nature of this responsivity.

Materials and Methods

In Experiment 1, whole-brain fMRI data were collected during a backward masking paradigm involving three levels of masked facial trustworthiness (low, average, high), adapted from previous studies using a blocked design to maximize statistical power (Whalen et al., 1998; Kim et al., 2010). In Experiment 2, we extended the backward masking paradigm to a rapid event-related design that allowed us to test amygdala responsivity to a wider and fully continuous range of facial trustworthiness, and to directly compare neural activity between subliminal and supraliminal presentations.

Subjects. Twenty-one volunteers (16 females) between the ages of 18 and 22 years [mean (M) = 18.75 years] participated in Experiment 1, and 16 volunteers (12 females) between the ages of 18 and 35 years (M = 21.80 years) participated in Experiment 2. Two subjects in Experiment 1 and one subject in Experiment 2 were excluded because of reported awareness of the subliminal stimuli. All subjects in Experiments 1 and 2 were right-handed, native English speakers, and had normal or corrected-to-normal vision and no history of neurological disorders or use of psychoactive medications. They received partial course credit or monetary compensation.

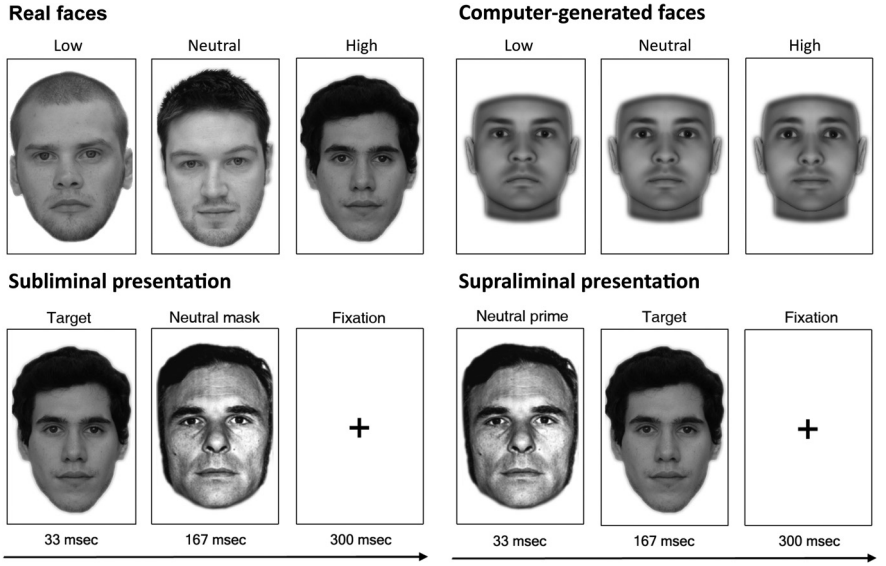


Figure 1. fMRI experimental procedure. Sample real stimuli (Experiments 1 and 2) and computer-generated stimuli (Experiment 1 only). For subliminal presentations (Experiments 1 and 2), the target face was presented for 33 ms and replaced by a neutral face mask for 167 ms that disrupted further visual processing of the target. This was followed by a 300 ms fixation period. For supraliminal presentations (Experiment 2), the target and neutral face were reversed.

Face stimuli. In a pretest, raters ($N = 10$) judged the trustworthiness of 300 neutral-affect, male and female faces from the Glasgow Unfamiliar Face Database (GUFD; Burton et al., 2010), normalized for size and luminance, in randomized order on a 7-point Likert scale. Judgments were highly consistent across the raters (intraclass correlation coefficient = 0.912). For computer-generated targets, we used well validated faces developed by previous research using 3D face modeling to generate faces varying in trustworthiness (Oosterhof and Todorov, 2008). There were 24 computer-generated faces in each of the 3 conditions (low, average, high), corresponding to −2 SD (low), M (average), and +2 SD (high) of the trustworthiness dimension developed in previous work (Oosterhof and Todorov, 2008). All facial targets in Experiments 1 and 2 were emotionally neutral. See Figure 1 for task design and sample stimuli.

In Experiment 1, half of the functional runs presented real facial targets from the GUFD, whereas the other half of functional runs presented computer-generated facial targets. For real targets, the pretest ratings were used to create 3 conditions: 24 low-trustworthy (M = 2.74, SE = 0.20), 24 average-trustworthy (M = 3.76, SE = 0.22), and 24 high-trustworthy (M = 4.67, SE = 0.21) faces. A mean rating per face, averaged across pretest raters, was computed. Low-trustworthy faces were rated as significantly less trustworthy than average-trustworthy faces ($t_{(46)} = 23.02$, $p < 0.0001$), which were rated significantly less trustworthy than high-trustworthy faces ($t_{(46)} = 19.02$, $p < 0.0001$). In Experiment 2, a total of 160 real male and female faces were used as the target stimuli, which were sampled evenly across the total distribution of 300 real faces varying in trustworthiness from Experiment 1 (stimuli from the GUFD).

In functional runs involving real faces, neutral-affect Ekman faces were used as mask stimuli. In functional runs involving computer-generated faces, separate computer-generated faces that were neutral-affect, invariant on trustworthiness, and developed by the same previous research as the target stimuli were used as mask stimuli.

Design of Experiment 1. Experiment 1 used a backward masking paradigm involving 3 levels of masked facial trustworthiness (low, average, high), adapted from previous backward-masking studies using a blocked design to maximize statistical power (Whalen et al., 1998; Kim et al., 2010). During fMRI, subjects viewed blocks of either low-, average-, or high-trustworthy facial targets that were masked by neutral face presentations. The targets in half of the functional scan runs were real faces from the GUFD, described above, which were prerated on trustworthiness, obtained using the same camera and lighting, and not differing in low-level visual properties. Targets in the other half of the scan runs were computer-generated faces developed in previous studies using statistical face modeling to convey either low, average, or high trustworthiness (Oosterhof and Todorov, 2008). The computer-generated faces varied only in trustworthiness cues; all other perceptual information was controlled.

In the scanner, subjects passively viewed blocks of the target stimuli and masks across 10 functional runs. The first five runs consisted of real facial targets, and the last five runs consisted of computer-generated facial targets. During each run, subjects viewed 9 blocks. The order of block sequences was counterbalanced across subjects. Target trustworthiness varied between blocks, with 3 blocks per trustworthiness condition (low, average, and high trustworthiness). Each block consisted of 24 target faces unique to that block, whose order was randomized within the block. Five masks were randomly paired with targets per block. The targets were each centered on a black background in isolation for a duration of 33 ms. A neutral face mask then immediately replaced the prime for 167 ms, after which an interstimulus interval of 300 ms (fixation cross) ensued (Fig. 1). To maintain subjects' visual attention, subjects participated in a 1-back task in which they were asked to press a button if the same face (mask) was presented twice. The blocks were interleaved by 12 s fixation-cross baseline blocks in each run. Following the last block, a 10 s fixation-cross period completed the run.

Analysis of Experiment 1. Individual subjects' BOLD signals in Experiment 1 were modeled using a general linear model (GLM) with 6 predictors: 3 (low-trustworthy, average-trustworthy, and high-trustworthy faces) × 2 (real faces, computer-generated faces). All predictors were modeled as boxcar functions across block durations and convolved with a two-gamma hemodynamic response function (HRF). First-level GLM analyses conducted on individual subjects' BOLD responses were submitted to a second-level random-effects analysis, treating subjects as a random factor.

Given an a priori hypothesis of amygdala involvement, we defined an anatomical region of interest (ROI) of the bilateral amygdala using the meta-analytic Neurosynth database (Yarkoni et al., 2011) as 10 mm spheres centered on [±18, −6, −11]. Parameter estimates (β values)

were extracted from the bilateral amygdala ROI and submitted to a 3 (trustworthiness: low, average, high trustworthiness) × 2 (stimulus type: real, computer-generated) repeated-measures ANOVA.

To further explore any additional effects of trustworthiness, we conducted random-effects analyses within a mask of the bilateral amygdala. Multiple statistical testing of voxels was corrected (false-positive rate $p < 0.05$) using a voxelwise threshold of $p < 0.05$ and a minimum cluster-size extent (k) of 421 mm^3. The minimum cluster-size extent needed to maintain an experiment-wide α of 0.05 was empirically determined by a Monte Carlo simulation, accounting for spatial correlations between neighboring voxels (Forman et al., 1995).

Design of Experiment 2. In Experiment 2, we extended the backward masking paradigm to a rapid event-related design that allowed us to test amygdala responsivity to a wider and fully continuous range of facial trustworthiness, and to directly compare neural activity between subliminal and supraliminal presentations. During fMRI, subjects took part in a modified event-related version of the backward masking paradigm using a considerably larger set of 160 target faces. Targets' trustworthiness varied trial-by-trial (each 2000 ms) under both subliminal and supraliminal conditions. Trustworthiness ratings were obtained for all faces postscan. Subjects passively viewed the target stimuli and neutral Ekman face masks across 4 functional runs. The first 2 runs presented targets subliminally, similar to Experiment 1. Targets were centered on a black background for 33 ms, which was immediately replaced by a neutral Ekman face mask for 167 ms. The mask was then followed by a fixation cross for 300 ms (Fig. 1). This 500 ms sequence repeated four consecutive times to maximize BOLD sensitivity to the targets and cover the length of a single TR (2000 ms). The second 2 runs involved supraliminal presentations, wherein the target faces and masks were reversed. Thus, a neutral Ekman face was centered on a black background for 33 ms, to be immediately followed by a target face for 167 ms and finally a 300 ms fixation cross. Like the first 2 subliminal runs, this sequence repeated four consecutive times (2000 ms). The design therefore ensured that visual information was identical across the subliminal versus supraliminal conditions. We refer to a single 2000 ms sequence as a "subliminal event" or "supraliminal event," respectively.

Of the total 160 faces used, two sets of 80 faces with matched trustworthiness distributions were then created. Each set was divided into 4 levels of 20 faces (low, medium-low, medium-high, high) to generate a number of presentation orders within a run that maximized statistical power for detecting parametric effects (i.e., to spread the trustworthiness variability across the run). For the first subliminal run, the 80 faces from one of the two sets (each face repeated twice) and an additional 20 baseline events (2000 ms of fixation cross) were used; for the second subliminal run, the 80 faces of the remaining set (each face repeated twice) and an additional 20 baseline events were used. The 2 supraliminal runs were identical to the 2 subliminal runs with the exception of reversing the target and mask stimuli. All events within runs were sequenced in a manner to optimize the efficiency of event-related BOLD signal estimation (Dale, 1999). Which face set was presented first versus second and the presentation order within runs were counterbalanced across subjects. A 6 s fixation-cross period divided the first and second half of each run (serving as a break), and another 6 s fixation-cross period completed each run. Note that it was not possible to counterbalance the order of subliminal versus supraliminal runs, as having supraliminal runs precede subliminal runs would heighten subjects' awareness of the target stimuli before their subliminal presentation.

After the scan, subjects were presented with each of the target faces one at a time in randomized order and rated their trustworthiness from 1 ("not at all") to 7 ("very much") using the keyboard. These ratings were used on a subject-by-subject, face-by-face basis to model BOLD responses.

Analysis of Experiment 2. In Experiment 2, subjects' postscan trustworthiness ratings were z-normalized. Individual subjects' BOLD signals were modeled using a GLM with 6 total predictors: 2 dichotomous predictors for modeling the presentation of subliminal and supraliminal events, 2 parametric predictors modeling a subject's postscan trustworthiness rating of the target for subliminal and supraliminal events (linear effects), and 2 parametric predictors modeling the square of a subject's trustworthiness rating of the target for subliminal and supraliminal events (quadratic effects). All predictors were modeled as boxcar functions (the amplitude of which was parametrically varied, for parametric predictors) and convolved with a two-gamma HRF. First-level GLM analyses conducted on individual subjects' fMRI signal were submitted to a second-level random-effects analysis, treating subjects as a random factor.

Using the anatomical ROI of the bilateral amygdala (see details in Analysis of Experiment 1), β values were extracted to test for linear and quadratic trustworthiness effects. To identify clusters of activation within the amygdala exhibiting significant parametric effects, we conducted random-effects parametric analyses using a restricted mask of the bilateral amygdala. We controlled for multiple statistical testing of voxels within the bilateral amygdala mask (false-positive rate $p < 0.05$) using the same correction technique as in Experiment 1.

To better specify the nature of the quadratic modulation in the amygdala, β values associated with the individual 160 face stimuli were extracted for each subject. These were submitted to polynomial regression analysis using a multilevel generalized estimating equation (GEE) approach that can incorporate such nested, trial-by-trial data while accounting for the intracorrelations due to repeated measurements (Zeger and Liang, 1986). In our case, this included intracorrelations associated with individual subjects and with individual face stimuli. Separately for the left and right amygdala, β values associated with the 160 targets were regressed onto linear and quadratic components of subjects' postscan trustworthiness ratings.

Postscan discrimination tasks. To provide an objective measure of awareness in Experiments 1 and 2, we used postscan discrimination tasks. While still in the scanner, subjects were presented with facial targets one at a time in randomized order using a procedure virtually identical to the one used in the main fMRI tasks. Before starting the discrimination task, subjects were informed of target face presentations occurring before the masks, and were asked to decide their gender. The accuracy of gender categorizations served as the measure of objective awareness. Each trial began with a 500 ms fixation cross, followed by the target face. After 33 ms, the target was immediately replaced by the mask for 167 ms. In the Experiment 1 discrimination task, the mask was followed by a prompt for gender categorization (male or female?). In the Experiment 2 discrimination task, the fixation cross, target face, and mask repeated four consecutive times (as was done in the Experiment 2 fMRI task) before the categorization prompt appeared. Subjects were instructed to categorize the gender of the target as quickly and accurately as possible using a button press; they were given unlimited time to make their response. Differing from the main fMRI tasks, trials were self-paced. Subjects completed several practice trials before starting the task. In the Experiment 1 discrimination task, subjects completed 144 trials (72 real face targets and 72 computer-generated face targets). Because Experiment 1 involved only male targets, we used a representative half of the faces used in the main fMRI task in addition to an equal number of filler female faces matched for visual properties. In the Experiment 2 discrimination task, subjects completed 160 trials (the 160 real faces—half male, half female—used in the main fMRI task). The same Ekman and computer-generated mask stimuli were used for real and computer-generated face targets, respectively.

Discrimination of facial trustworthiness. Although our postscan discrimination tasks in Experiments 1 and 2 assessed perceptions of gender to ensure a lack of discriminability for masked targets' facial characteristics in general, trustworthiness could potentially have remained discriminable despite gender being indiscriminable with our masking procedure. To alleviate this concern, an additional behavioral experiment was conducted in which 16 volunteers (10 females) between the ages of 18 and 24 years (M = 19.19 years) participated. The procedure was identical to the postscan discrimination tasks of Experiments 1 and 2, except that participants were prompted to judge the target's trustworthiness (untrustworthy or trustworthy?) rather than gender. Trials were self-paced and subjects were given unlimited time to make their response. After completing several practice trials, subjects were presented with 211 trials, comprising the 72 non-filler targets used in the Experiment 1 discrimination task and 160 targets used in the Experiment 2

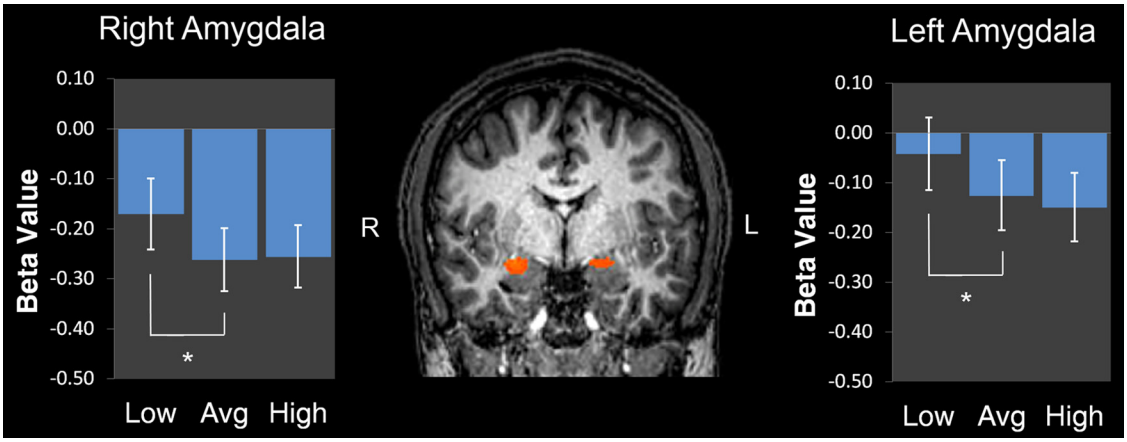


Figure 2. Stronger bilateral amygdala activation to low-trustworthy faces. Coronal slice ($y = -5$), depicting stronger responses to low-trustworthy targets versus average- or high-trustworthy targets from a random-effects analysis targeting the bilateral amygdala ($p < 0.05$, corrected). Bar plots depict mean β values for the 3 block types. Error bars indicate SEM.

discrimination task (there were 21 targets that overlapped in both experiments, resulting in a total of 211 trials). Following this initial phase of the experiment, subjects viewed all targets again in a new randomized order, one at a time, with unlimited exposure (no masking), and they were asked to judge trustworthiness along a 6-point Likert scale. We used a 6-point rather than a 7-point scale to be able to dichotomize judgments (0–3 = untrustworthy; 4–6 = trustworthy), thereby permitting signal detection analysis and a controlling of response bias.

fMRI acquisition and preprocessing. In both experiments, subjects were scanned using a 3T Philips Intera Achieva Scanner (Philips Medical Systems) equipped with a SENSE birdcage head coil in the Dartmouth Brain Imaging Center. All stimuli were back-projected onto a screen visible via a mirror mounted on the MRI head coil (visual angle ~13.5 × 13.5°). Anatomical images were acquired using a T1-weighted protocol (256 × 256 matrix, 128 1.33 mm transverse slices). Functional images were acquired using a single-shot gradient echo EPI sequence (TR = 2000 ms, TE = 35 ms). Thirty-five interleaved oblique-axial slices (3 mm × 3 mm × 4 mm voxels; no slice gap) parallel to the AC-PC line were obtained. Analysis of the imaging data was conducted using BrainVoyagerQX (Brain Innovation). Functional imaging data preprocessing included 3D motion correction, slice-timing correction (sinc interpolation), spatial smoothing using a 3D Gaussian filter (8 mm FWHM), and voxelwise linear detrending and high-pass filtering of frequencies (>3 cycles per time course). Structural and functional data of each subject were transformed to standard Talairach stereotaxic space.

Results

Experiment 1

Experiment 1 presented subjects with subliminally presented faces that varied on three levels of trustworthiness (low, average, high). Subjects were questioned about their subjective awareness of the subliminal stimuli after the scan; two subjects reported subjective awareness and were excluded. Postscan discrimination data from the remaining subjects were analyzed using signal detection to appropriately control for response bias; signal was arbitrarily defined as female. Thus, perceptual discriminability (d') was computed as the percentage of masked female primes that were successfully categorized as female (hits), adjusted for the percentage of male masked primes that were erroneously categorized as female (false alarms): d' = z-score (% hits) − z-score (% false alarms), with chance performance = 0 ± 1.74. No included subject's discriminability (d') rose significantly above chance (all p values <1.74), and d' overall was quite low (M = 0.17, SE = 0.11), ensuring that the masked stimuli were below subjects' awareness.

Given an a priori hypothesis of amygdala responsivity, we examined activation in the bilateral amygdala using anatomically defined, 10 mm spherical ROIs. A repeated-measures ANOVA conducted on parameter estimates (β values) extracted from the bilateral amygdala ROI revealed a significant effect of facial trustworthiness, $F_{(2,36)} = 3.80$, $p = 0.032$. Paired-samples t tests indicated that this effect was driven specifically by low levels of trustworthiness. Low-trustworthy targets elicited stronger activation than average-trustworthy targets ($t_{(18)} = 2.12$, $p = 0.048$); however, average- and high-trustworthy targets were not distinguished ($t_{(18)} = 0.54$, $p = 0.593$; Fig. 2*B*). The main effect of stimulus type (real vs computer-generated; $F_{(1,18)} = 0.40$, $p = 0.534$) and the interaction ($F_{(2,36)} = 0.33$, $p = 0.720$) were not significant; further analyses therefore were collapsed across real and computer-generated facial targets. To ensure that any possible weak, albeit nonsignificant, discriminability (d') of the masked stimuli was not driving this effect, subjects' d' was included as a covariate. However, responsivity in the bilateral amygdala to trustworthiness persisted when controlling for d' ($F_{(2,34)} = 3.477$, $p = 0.042$), indicating that discriminability had a negligible impact on the effect. Furthermore, the interaction between trustworthiness and d' was not significant ($F_{(2,34)} = 0.069$, $p = 0.933$), alleviating the concern that any potential visibility confounded the amygdala responses to trustworthiness.

To examine the possibility of additional trustworthiness effects, a random-effects ANOVA tested for clusters of activation that exhibited modulation by trustworthiness using a restricted mask of the bilateral amygdala. This revealed clusters in the right ($x = 22, y = -4, z = -10$, 429 mm^3, $F = 4.11$) and left ($x = -21$, $y = -9$, $z = -9$, 769 mm^3, $F = 3.81$) amygdala that showed especially strong responses to low-trustworthy targets ($p < 0.05$, corrected; Fig. 2), corroborating the anatomical ROI analysis.

To specifically identify any amygdala regions that were responsive to both low and high trustworthiness (a nonlinear effect), we conducted a random-effects contrast of [low + high] > average. This revealed a cluster in the right amygdala located more posteriorly ($x = 21, y = -12, z = -10$, 677 mm^3, $F = 2.32$; $p < 0.05$, corrected). The β values were extracted from this region, and a within-subject ANOVA contrast [1, −2, 1] confirmed a quadratic response, $F_{(1,18)} = 9.380$, $p = 0.007$. Planned comparisons indicated that, relative to average-trustworthy faces, both low-trustworthy ($t_{(18)} = 2.70$, $p = 0.015$) and high-trustworthy ($t_{(18)} = 2.11$, $p = 0.049$) faces more strongly engaged this region (one-tailed tests for directional hypothesis; Fig. 3). Similar to the negative-linear effect in the bilateral amygdala, the

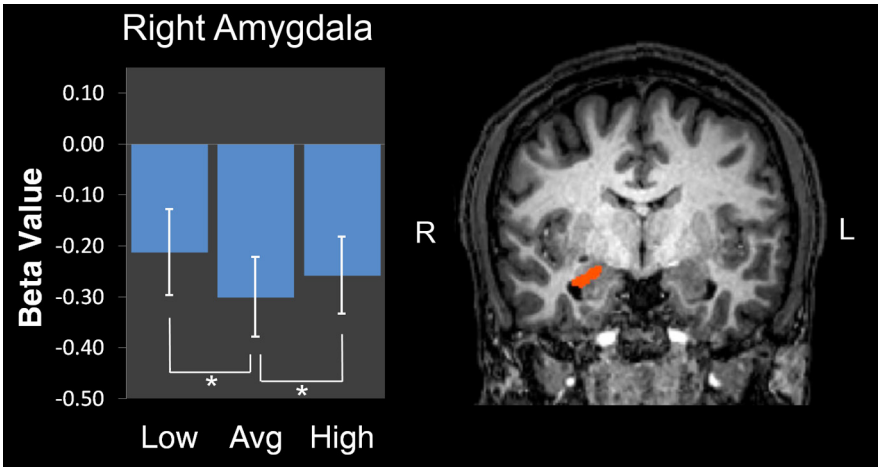


Figure 3. Region of the right amygdala responsive to both low and high levels of trustworthiness. Coronal slice ($y = -7$) depicting a posterior region of the right amygdala that responded more strongly to low- and high-trustworthy faces than average-trustworthy faces from a random-effects analysis targeting the bilateral amygdala ($p < 0.05$, corrected). Bar plots depict mean β values for the 3 block types. Error bars indicate SEM.

quadratic effect in this posterior region of the right amygdala remained significant after including d' as a covariate ($F_{(1,17)} = 7.642$, $p = 0.013$); further, the interaction between trustworthiness and d' was not significant ($F_{(1,17)} = 0.002$, $p = 0.964$), thus ensuring that any potential visibility did not confound this result.

Given the critical role of the fusiform cortex in face perception (Haxby et al., 2000), exploratory random-effects analyses tested for such linear and nonlinear effects of trustworthiness within a mask of the bilateral fusiform cortex; however, no effects were present that survived correction.

Thus, regions in the bilateral amygdala exhibited especially strong activation for low-trustworthy faces when presented subliminally. Moreover, a region of the posterior right amygdala exhibited a nonlinear pattern of response, showing stronger activation to both low and high trustworthiness. These findings extend previous reports of linear and nonlinear amygdala responses to trustworthiness during supraliminal presentations to the subliminal level.

Interestingly, amygdala activation was reduced below baseline across the three trustworthiness conditions. This may be because of deactivations of the amygdala commonly observed in tasks required high-level cognitive processing. Specifically, because experimental blocks involved a 1-back task used to maintain subjects' attention, the demands of the 1-back task may have overall reduced amygdala responses relative to baseline blocks that required the mere passive viewing of a fixation cross. Previous work suggests that cognitively demanding tasks (e.g., *n*-back tasks) suppress amygdala response to visual stimuli (Drevets and Raichle, 1998), even when those stimuli are task-relevant (Kellermann et al., 2012). Thus, one explanation for reduced amygdala responses during experimental blocks is the additional cognitive demands associated with those blocks. Despite an overall reduced response in the amygdala, however, the primary result and the amygdala's relative response to subliminal variation in facial trustworthiness is clear.

In Experiment 2, we expanded on these results by testing amygdala responsivity to a full continuum of trustworthiness using a rapid event-related design, and directly comparing with supraliminal presentations.

Experiment 2

After the scan, one subject reported subjective awareness of the targets and was excluded. As in Experiment 1, postscan forced-choice discrimination data were analyzed using signal detection. No subject's discriminability (d') rose significantly above chance (all p values <1.74), and overall d' was even negative ($M = -0.09$, $SE = 0.06$), ensuring that the masked stimuli were below subjects' awareness.

Parametric analyses simultaneously modeled linear and nonlinear (quadratic) predictors based on subjects' postscan trustworthiness ratings. Activation averaged across the anatomical ROI of the bilateral amygdala was significantly modulated by a quadratic trustworthiness effect ($t_{(14)} = 2.25$, $p = 0.041$); no linear effect emerged ($t_{(14)} = 0.30$, $p = 0.769$). To investigate further, a random-effects parametric analysis targeting the bilateral amygdala was probed for clusters exhibiting significant quadratic responses ($p < 0.05$, corrected). This revealed clusters in the left ($x = -22$, $y = -5$, $z = -17$; $k = 2495$ mm^3; $t = 3.64$) and right ($x = 19$, $y = -6$, $z = -16$; $k = 1559$ mm^3; $t = 3.51$) amygdala (Fig. 4, top), consistent with the anatomical ROI analysis. The β values were extracted to examine the parametric effects separately for subliminal and supraliminal conditions. The quadratic effect was significant during supraliminal presentations (left: $t_{(14)} = 2.17$, $p = 0.048$; right: $t_{(14)} = 2.55$, $p = 0.023$]. Critically, it was also significant during subliminal presentations (left: $t_{(14)} = 2.17$, $p = 0.048$; right: $t_{(14)} = 2.35$, $p = 0.034$). Linear effects were not significant in these clusters for either supraliminal (left: $t_{(14)} = -0.15$, $p = 0.885$; right: $t_{(14)} = -1.19$, $p = 0.252$) or subliminal (left: $t_{(14)} = 1.81$, $p = 0.091$; right: $t_{(14)} = 0.25$, $p = 0.804$) conditions (Fig. 4, middle). A random-effects parametric analysis testing for significant linear responses revealed no clusters of activation that survived correction.

To better specify the nature of the quadratic modulation, an additional model was constructed to extract β values uniquely associated with the individual 160 face stimuli. Extracted β values were submitted to polynomial regression analyses using a GEE approach that can incorporate such nested, trial-by-trial data while accounting for intracorrelations due to repeated measurements (Zeger and Liang, 1986). In both the left and right amygdala, the linear effect of trustworthiness was not significant (left: $B = 0.010$, $Z = 0.38$, $p = 0.705$; right: $B = 0.002$, $Z = 0.06$, $p = 0.953$), whereas the quadratic effect was highly significant (left: $B = 0.040$, $Z = 2.66$, $p = 0.008$; right: $B = 0.037$, $Z = 2.44$, $p = 0.015$). This quadratic effect did not interact with presentation condition (subliminal vs supraliminal; left: $B = 0.0166$, $Z = 0.61$, $p = 0.540$; right: $B = 0.0319$, $Z = 1.12$, $p = 0.262$), and thus, further analyses collapsed across them. Plotting mean β values in the left and right amygdala revealed a U-shaped encoding function. For both supraliminal and subliminal presentations, the amygdala increased in activation as faces approached the low and high ends of the continuum, relative to average faces at the middle (Fig. 4, middle and bottom). As in Experiment 1, d' was included as a covariate to control for any possible influence of weak discriminability of masked stimuli. The quadratic effect remained significant after including d' as a covariate (left: $B =$

0.0416, $Z = 2.77$, $p = 0.006$; right: $B = 0.0402$, $Z = 2.66$, $p = 0.008$) and did not significantly interact with d' (left: $B = -0.0734$, $Z = -0.92$, $p = 0.358$; right: $B = -0.1122$, $Z = -1.43$, $p = 0.152$), eliminating concerns of possible visibility confounding the effect.

Finally, as in Experiment 1, exploratory random-effects analyses tested for any linear and nonlinear responses within a mask of the bilateral fusiform cortex; however, no effects were present that survived correction.

Thus, the results of Experiment 2 replicate the nonlinear effects of trustworthiness obtained in previous studies using supraliminal presentations, and show that such effects extend to subliminal exposures outside subjects' awareness.

We should note that, because the amygdala can exhibit adaptation to emotionally significant stimuli in masking paradigms (Whalen et al., 1998), it is possible that the visible repetition of a small number of Ekman masks in the subliminal condition may have elicited adaptation effects not present in the supraliminal condition (which visibly presented 160 targets without repetition). This may be additionally important because amygdala responses can be shaped by the conjunction of masked and masking stimuli (Kim et al., 2010). However, because the nonlinear amygdala response to trustworthiness was statistically identical across the subliminal and supraliminal conditions, this is unlikely to have been an issue in the present experiment.

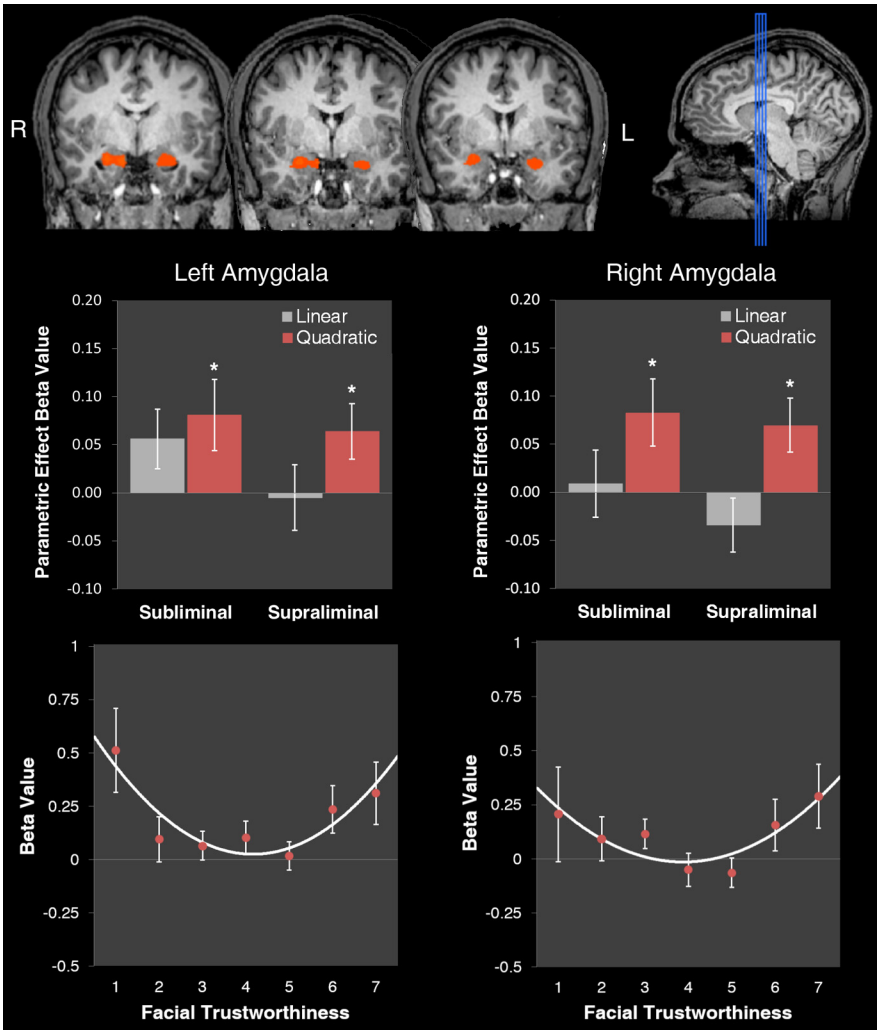


Figure 4. Nonlinear amygdala responses to facial trustworthiness during subliminal and supraliminal presentations. Top, Coronal slices ($y = -1 \pm 2$) depicting responses associated with a quadratic trustworthiness effect from a random-effects analysis targeting the bilateral amygdala ($p < 0.05$, corrected). Middle, Mean β values for the linear and quadratic parametric predictors, indexing the strength of the correlations between amygdala response and linear vs quadratic components of trustworthiness, separately for left amygdala (left) and right amygdala (right). Bottom, Mean β values for amygdala BOLD response as a function of facial trustworthiness (for both supraliminal and subliminal conditions), separately for left amygdala (left) and right amygdala (right). Error bars indicate SEM.

Discrimination of masked trustworthiness

An additional behavioral experiment was run to ensure that facial trustworthiness in particular was not discriminable from the masking procedure we used (see Materials and Methods). No subject reported subjective awareness of the targets after the experiment. Forced-choice discrimination data (0 = untrustworthy; 1 = trustworthy) were analyzed using signal detection. A face's "correct" trustworthiness was defined by a subject's own rating of that specific face following the discrimination task, in which a 6-point Likert scale was used to be able to dichotomize ratings (1–3 = untrustworthy; 4–6 = trustworthy). No subject's d' rose significantly above chance (all p values <1.74), and overall d' was quite small (M = 0.03, SE = 0.06), ensuring that the trustworthiness of masked faces was not discriminable.

Correspondingly, we should note that although subjects did not report any subjective awareness of masked faces, it is possible a more objective measure of awareness would have shown evidence of subjects detecting some masked stimuli as faces (as opposed to non-faces). Indeed, the basic detection of a face is dissociable from the discrimination of facial characteristics such as trustworthiness or gender. For instance, one recent study found evidence for accurate discrimination of facial emotion despite the facial stimuli being so impoverished they were explicitly categorized as shoes (Seirafi et al., 2013). Nevertheless, what is critical is that facial trustworthiness in particular was not discriminable in the subliminal presentations we used, as indicated by the aforementioned analysis. This specifically shows that the amygdala can track high-level social information from a face (e.g., trustworthiness) without that information being consciously perceived.

Discussion

Across two experiments, our findings demonstrate that the human amygdala is automatically responsive to a face's trustworthiness in the absence of perceptual awareness. The amygdala has been shown to code trustworthiness in an attention-independent fashion when faces are clearly visible (Engell et al., 2007; Todorov et al., 2011). However, the notion that a high-level facial characteristic such as trustworthiness may be assessed before a face is consciously perceived has only been recently suggested by behavioral studies (Todorov et al., 2009;

Stewart et al., 2012), and the role of the amygdala in this process has remained unclear.

The present results are striking in that they raise a conundrum as to how the amygdala could evaluate such complex social information under constricted processing. Researchers have documented amygdala responses to subliminal fear expressions (Whalen et al., 1998), which have been suggested by some to be related to an early threat-detection mechanism driven by a retino-collicular-pulvinar pathway that responds to salient, behaviorally relevant stimuli (Morris et al., 1999; Davis and Whalen, 2001; LeDoux, 2007). However, amygdala responses to subliminal fearful expressions may be driven by simple visual cues such as enlarged eye whites (Whalen et al., 2004). In contrast, the features associated with facial trustworthiness are considerably more complex (Oosterhof and Todorov, 2008), although at extreme levels, they do overlap with emotional expressions of anger and joy (Oosterhof and Todorov, 2009), and some studies have found amygdala activation to masked angry faces (Morris et al., 1998). Nonconscious processing of faces has been reported within ventral temporal cortex alone (Green et al., 2005; Sterzer et al., 2009), where detailed facial information can be extracted (Haxby et al., 2000; Kanwisher and Yovel, 2006). Limited processing via the ventral temporal cortex could potentially drive amygdala activation to subliminal stimuli as well, and thus there are multiple routes through which facial trustworthiness could affect amygdala responsivity without perceptual awareness (Pessoa and Adolphs, 2010; Tamietto and de Gelder, 2010). Such ventral-temporal processing during masked exposures has been shown to exhibit face-specific encoding (Sterzer et al., 2009), which could potentially contribute to amygdala responsivity (Pessoa and Adolphs, 2010; Tamietto and de Gelder, 2010). Indeed, facial trustworthiness is processed extensively by face-specific areas in the ventral-visual stream, such as the fusiform cortex (Said et al., 2010; Mende-Siedlecki et al., 2013; Todorov et al., 2013).

It is also quite possible that both cortically and subcortically driven routes to the amygdala could be involved in the subliminal effects, with the critical question being the degree of cortico-subcortical interaction. In this perspective, it is the feedback loops between subcortical and cortical areas that are thought to be critical for awareness, rather than the involvement of particular brain regions. Thus, masked exposures to facial trustworthiness may have been sufficient to modulate amygdala responses because of enough bottom-up information arriving through feedforward connections (whether they be subcortical or cortical), but not sufficient to engage the more elaborate reentrant cortico-subcortical feedback important for awareness (Williams et al., 2006; Pessoa and Adolphs, 2010). As such, masked trustworthiness could have activated limited processing in any number of regions that possibly contributed to amygdala responses here, but it may not have been able to be consciously perceived due to a lack of cortico-subcortical interaction. Future research could clarify the specific pathways that underlie the amygdala's tracking of trustworthiness outside awareness.

If, however, ventral-stream face processing participated in the amygdala responsivity here, one would have expected face-processing areas (e.g., fusiform cortex) to be modulated by trustworthiness. In the present experiments, the fusiform cortex was not reliably modulated by trustworthiness, consistent with some previous studies finding robust amygdala effects of trustworthiness without significant fusiform effects (Engell et al., 2007). One possibility is that our current methodological approach was not sensitive enough to detect trustworthiness effects in the fusiform cortex. For instance, because the fusiform cortex is a critical face-processing region, it exhibits particularly strong adaptation to visual facial features (Kanwisher and Yovel, 2006). Because of the nature of our masking procedure and because the target and mask stimuli were both emotionally neutral faces and highly similar, it is possible that fusiform responses were confounded by adaptation effects between target and mask stimuli. If these effects occurred, they would have diminished sensitivity to trustworthiness responses in the fusiform cortex considerably. Future work might therefore consider using different masking techniques to focus on fusiform responsivity. There is also the possibility of a genuine dissociation between amygdala and fusiform modulation by trustworthiness, such that the amygdala exhibits responses to subliminally presented trustworthiness without the fusiform exhibiting such responses. If this were true, however, reliable fusiform modulation would have been observed during the supraliminal condition, which was not the case. Further studies that target fusiform responsivity directly could address these possibilities.

It should also be noted that although trustworthiness is regarded as a fundamental dimension of social evaluation (Oosterhof and Todorov, 2008), clearly, it is composed of lower-level featural dimensions that jointly contribute to its perception. For example, features on an emotionally neutral face exhibiting subtle structural overlap with particular emotion expressions (e.g., anger or joy) contribute to perceived trustworthiness (Said et al., 2009). Complex arrangements of such lower-level facial features together form this fundamental trustworthiness dimension, which is spontaneously perceived on encountering a face and accounts for the majority of variance in social evaluation (Engell et al., 2007; Stewart et al., 2012). Here we show that this dimension can sensitively modulate amygdala responses before reaching awareness.

It is only recently that researchers have begun to examine linear and nonlinear responses to facial trustworthiness in tandem (Said et al., 2010; Todorov et al., 2011; Mende-Siedlecki et al., 2013). In some studies, as in Experiment 1, such effects cohabitate within different subregions of the amygdala (Todorov et al., 2013); in other studies, as in Experiment 2, quadratic effects trump linear effects entirely (Said et al., 2010). Work in nonhuman primates has also converged on the finding of both forms of coding in the amygdala, with one subregion showing linear responses to threatening faces (coding negative valence), and another showing nonlinear responses to both threatening and appeasing faces relative to neutral (coding salience; Hoffman et al., 2007). In humans, neuroimaging work more generally has dissociated amygdala subregions coding for the valence versus salience of stimuli in ambiguous learning situations, e.g., a novel social encounter (Whalen, 1998; Kim et al., 2003; Davis et al., 2010). The results of Experiment 1 are thus consistent with this prior research, finding coexisting linear and nonlinear representations; however, Experiment 2 obtained evidence only for a nonlinear representation. Although also consistent with prior work, the difference could be driven by the nature of the blocked versus event-related tasks. Repeated presentations in the blocked design of Experiment 1 were similar to the task design of Kim et al. (2003) and may have induced a task context that increased the importance of tracking valence over salience. Alternatively, the event-related design of Experiment 2 used a wider, continuous range of trustworthiness, which may have increased sensitivity to nonlinear effects (Todorov et al., 2013). Despite these differences between blocked versus event-related presentations, both experiments provide clear support for our primary hypothesis that the

amygdala is automatically responsive to facial trustworthiness without perceptual awareness. Regions in the amygdala track how untrustworthy an unseen face appears as well as the overall strength of the trustworthiness signal. Moreover, this tracking generalizes across real and computer-generated faces, where trustworthiness was both measured and manipulated, respectively.

Whereas the negative-linear effects found here are consistent with the amygdala's role in vigilance for threats and tracking of valence (Davis and Whalen, 2001; LeDoux, 2007), nonlinear effects are consistent with the amygdala's processing of salience (Zald, 2003) and motivational relevance (Phelps and LeDoux, 2005; Cunningham et al., 2008; Adolphs, 2010). Faces with stronger cues for untrustworthiness or trustworthiness are more motivationally relevant, as these cues spontaneously elicit approach and avoidance behaviors (Slepian et al., 2012). Both forms of coding are consistent with the amygdala's importance for interpreting implicit social signals (Adolphs et al., 1998; Freeman et al., 2010; Heberlein and Adolphs, 2004). Faces that appear more untrustworthy and likely to inflict harm, or faces with a stronger trustworthiness signal in general, would benefit more from automatic amygdala responsivity, which could adaptively modulate cortical processes and motivate appropriate social behavior (Davis and Whalen, 2001; Phelps and LeDoux, 2005).

In summary, we demonstrated in two experiments that the amygdala is sensitive to subliminal variation in facial trustworthiness. Thus, the amygdala can be influenced by even high-level facial information before that information is consciously perceived. These findings provide evidence that the amygdala's processing of social cues in the absence of awareness may be more extensive than previously described.

References

Adolphs R (2010) What does the amygdala contribute to social cognition? Ann N Y Acad Sci 1191:42–61. CrossRef Medline

Adolphs R, Tranel D, Damasio AR (1998) The human amygdala in social judgment. Nature 393:470–474. CrossRef Medline

Bar M, Neta M, Linz H (2006) Very first impressions. Emotion 6:269–278. CrossRef Medline

Burton AM, White D, McNeill A (2010) The Glasgow face matching test. Behav Res Methods 42:286–291. CrossRef Medline

Chagnon NA (1988) Life histories, blood revenge, and warfare in a tribal population. Science 239:985–992. CrossRef Medline

Cunningham WA, Van Bavel JJ, Johnsen IR (2008) Affective flexibility: evaluative processing goals shape amygdala activity. Psychol Sci 19:152–160. CrossRef Medline

Dale AM (1999) Optimal experimental design for event-related fMRI. Hum Brain Mapp 8:109–114. CrossRef Medline

Davis FC, Johnstone T, Mazzulla EC, Oler JA, Whalen PJ (2010) Regional response differences across the human amygdaloid complex during social conditioning. Cereb Cortex 20:612–621. CrossRef Medline

Davis M, Whalen PJ (2001) The amygdala: vigilance and emotion. Mol Psychiatry 6:13–34. CrossRef Medline

Drevets WC, Raichle ME (1998) Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive processes: implications for interactions between emotion and cognition. Cognition Emotion 12:353–385. CrossRef

Engell AD, Haxby JV, Todorov A (2007) Implicit trustworthiness decisions: automatic coding of face properties in the human amygdala. J Cogn Neurosci 19:1508–1519. CrossRef Medline

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med 33:636–647. CrossRef Medline

Freeman JB, Schiller D, Rule NO, Ambady N (2010) The neural origins of superficial and individuated judgments about ingroup and outgroup members. Hum Brain Mapp 31:150–159. CrossRef Medline

Green MF, Glahn D, Engel SA, Nuechterlein KH, Sabb F, Strojwas M, Cohen MS (2005) Regional brain activity associated with visual backward masking. J Cogn Neurosci 17:13–23. CrossRef Medline

Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. Trends Cogn Sci 4:223–233. CrossRef Medline

Heberlein AS, Adolphs R (2004) Impaired spontaneous anthropomorphizing despite intact perception and social knowledge. Proc Natl Acad Sci U S A 101:7487–7491. CrossRef Medline

Hoffman KL, Gothard KM, Schmid MC, Logothetis NK (2007) Facial-expression and gaze-selective responses in the monkey amygdala. Curr Biol 17:766–772. CrossRef Medline

Kanwisher N, Yovel G (2006) The fusiform face area: a cortical region specialized for the perception of faces. Philos Trans R Soc Lond B Biol Sci 361:2109–2128. CrossRef Medline

Kellermann TS, Sternkopf MA, Schneider F, Habel U, Turetsky BI, Zilles K, Eickhoff SB (2012) Modulating the processing of emotional stimuli by cognitive demand. Soc Cogn Affect Neurosci 7:263–273. CrossRef Medline

Kim H, Somerville LH, Johnstone T, Alexander AL, Whalen PJ (2003) Inverse amygdala and medial prefrontal cortex responses to surprised faces. Neuroreport 14:2317–2322. CrossRef Medline

Kim MJ, Loucks RA, Neta M, Davis FC, Oler JA, Mazzulla EC, Whalen PJ (2010) Behind the mask: the influence of mask-type on amygdala response to fearful faces. Soc Cogn Affect Neurosci 5:363–368. CrossRef Medline

LeDoux JE (2007) The amygdala. Curr Biol 17:R868–R874. CrossRef Medline

McDonald MM, Navarrete CD, Van Vugt M (2012) Evolution and the psychology of intergroup conflict: the male warrior hypothesis. Philos Trans R Soc B Biol Sci 367:670–679. CrossRef

Mende-Siedlecki P, Said CP, Todorov A (2013) The social evaluation of faces: a meta-analysis of functional neuroimaging studies. Soc Cogn Affect Neurosci 8:285–299. CrossRef Medline

Morris JS, Ohman A, Dolan RJ (1998) Conscious and unconscious emotional learning in the human amygdala. Nature 393:467–470. CrossRef Medline

Morris JS, Ohman A, Dolan RJ (1999) A subcortical pathway to the right amygdala mediating "unseen" fear. Proc Natl Acad Sci U S A 96:1680–1685. CrossRef Medline

Öhman A (2005) The role of the amygdala in human fear: automatic detection of threat. Psychoneuroendocrinology 30:953–958. CrossRef Medline

Oosterhof NN, Todorov A (2008) The functional basis of face evaluation. Proc Natl Acad Sci U S A 105:11087–11092. CrossRef Medline

Oosterhof NN, Todorov A (2009) Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. Emotion 9:128–133. CrossRef Medline

Pessoa L, Adolphs R (2010) Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. Nat Rev Neurosci 11:773–783. CrossRef Medline

Phelps EA, LeDoux JE (2005) Contributions of the amygdala to emotion processing: from animal models to human behavior. Neuron 48:175–187. CrossRef Medline

Said CP, Sebe N, Todorov A (2009) Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. Emotion 9:260–264. CrossRef Medline

Said CP, Dotsch R, Todorov A (2010) The amygdala and FFA track both social and non-social face dimensions. Neuropsychologia 48:3596–3605. CrossRef Medline

Seirafi M, De Weerd P, de Gelder BL (2013) Emotion categorization does not depend on explicit face categorization. J Vis 13(2):12 1–9. CrossRef

Slepian ML, Young SG, Rule NO, Weisbuch M, Ambady N (2012) Embodied impression formation: social judgments and motor cues to approach and avoidance. Soc Cogn 30:232–240. CrossRef

Sterzer P, Jalkanen L, Rees G (2009) Electromagnetic responses to invisible face stimuli during binocular suppression. Neuroimage 46:803–808. CrossRef Medline

Stewart LH, Ajina S, Getov S, Bahrami B, Todorov A, Rees G (2012) Unconscious evaluation of faces on social dimensions. J Exp Psychol Gen 141:715–727. CrossRef Medline

Tamietto M, de Gelder B (2010) Neural bases of the non-conscious perception of emotional signals. Nat Rev Neurosci 11:697–709. CrossRef Medline

Todorov A, Engell AD (2008) The role of the amygdala in implicit evalua-

tion of emotionally neutral faces. Soc Cogn Affect Neurosci 3:303–312. CrossRef Medline

Todorov A, Pakrashi M, Oosterhof NN (2009) Evaluating faces on trustworthiness after minimal time exposure. Soc Cogn 27:813–833. CrossRef

Todorov A, Said CP, Oosterhof NN, Engell AD (2011) Task-invariant brain responses to the social value of faces. J Cogn Neurosci 23:2766–2781. CrossRef Medline

Todorov A, Mende-Siedlecki P, Dotsch R (2013) Social judgments from faces. Curr Opin Neurobiol 23:373–380. CrossRef Medline

Whalen PJ (1998) Fear, vigilance, and ambiguity: initial neuroimaging studies of the human amygdala. Curr Dir Psychol Sci 7:177–188. CrossRef

Whalen PJ, Rauch SL, Etcoff NL, McInerney SC, Lee MB, Jenike MA (1998) Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. J Neurosci 18:411–418. Medline

Whalen PJ, Kagan J, Cook RG, Davis FC, Kim H, Polis S, McLaren DG, Somerville LH, McLean AA, Maxwell JS, Johnstone T (2004) Human amygdala responsivity to masked fearful eye whites. Science 306:2061. CrossRef Medline

Williams LM, Das P, Liddell BJ, Kemp AH, Rennie CJ, Gordon E (2006) Mode of functional connectivity in amygdala pathways dissociates level of awareness for signals of fear. J Neurosci 26:9264–9271. CrossRef Medline

Willis J, Todorov A (2006) First impressions: making up your mind after a 100-ms exposure to a face. Psychol Sci 17:592–598. CrossRef Medline

Winston JS, Strange BA, O'Doherty J, Dolan RJ (2002) Automatic and intentional brain responses during evaluation of trustworthiness of faces. Nat Neurosci 5:277–283. CrossRef Medline

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011) Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 8:665–670. CrossRef Medline

Zald DH (2003) The human amygdala and the emotional evaluation of sensory stimuli. Brain Res Rev 41:88–123. CrossRef Medline

Zebrowitz LA, Montepare JM (2008) Social psychological face perception: why appearance matters. Soc Personal Psychol Compass 2:1497. CrossRef Medline

Zeger SL, Liang KY (1986) Longitudinal data analysis for discrete and continuous outcomes. Biometrics 42:121–130. CrossRef Medline

Children show increase in mental health difficulties over COVID-19 lockdown

16 June 2020

	Emotional difficulties	Behavioural difficulties	Restlessness & attention difficulties
Children (parent report)	Increased	Increased	Increased
Adolescents (parent report)	Decreased	No change	Increased
Adolescents (adolescent report)	No change	No change	No change
Children with SEN & mental health difficulties	Decreased	No change	No change

Credit: University of Oxford

Parents/carers of children aged 4-10 years of age reported that over a one-month period in lockdown, they saw increases in their child's emotional difficulties, such as feeling unhappy, worried, being clingy and experiencing physical symptoms associated with worry, according to early results from the Co-SPACE study, asking parents and carers about their children's mental health through the COVID-19 crisis.

Over a one-month period in lockdown:

- Parents/carers of primary school age children taking part in the survey report an increase in their child's emotional, behavioral, and restless/attentional difficulties.
- Parents/carers of secondary school age children report a reduction in their child's emotional difficulties, but an increase in restless/attentional behaviors.
- Adolescents taking part in the survey report no change in their own emotional or behavioral, and restless/attentional difficulties.
- Parents/carers of children with Special Educational Needs (SEN) and those with a pre-existing mental health difficulty report a reduction in their child's emotional difficulties and no change in behavioral or restless/attentional difficulties.

More than 10,000 parents have now taken part in the Co-SPACE (COVID-19 Supporting Parents, Adolescents, and Children in Epidemics) survey led by experts at the University of Oxford.

Parents/carers also reported that their children's behavior had got worse over time, with an increase in behaviors such as temper tantrums, arguments and children not doing what they are asked. Parents/carers in the survey also reported that their children showed greater levels of restlessness/fidgety behavior and difficulties concentrating over the one month period.

Perhaps surprisingly, the same pattern was not seen in the older age group of 11-16 year olds. Teenagers themselves reported no change in their emotional difficulties between the two time points and their parents/carers reported that they felt that their child's emotional difficulties had actually improved. Neither teenagers nor their parents reported any changes in their behavior over this time but parents felt that their children were more restless and had more difficulty concentrating over time.

Tom Madders, Campaigns Director at YoungMinds, said, 'This research suggests that many younger children have found it increasingly hard to cope as the lockdown period has gone on, which may be because of loneliness, fears about the coronavirus or a loss of the routines and support that come with school.

'The picture appears to be more variable for older children in this study. Following the anxiety and uncertainty of going into lockdown, some are likely

to have found the restrictions more difficult as time as gone on, while others—including those who feel safe and secure at home but who find school challenging—may have adapted well to their new reality. For those young people, going back to school after a long break may well be tough, and it's vital that there's a re-adjustment period where wellbeing is prioritized.

'It's also important to recognize that some of the most vulnerable young people in our society—including those who have experienced abuse, violence or neglect—are often the hardest to identify. We need to ensure that effective support is available for all children who need it now and as restrictions lift.'

Professor Cathy Creswell, Professor of Developmental Clinical Psychology, University of Oxford, and co-leading the study, said, 'Prioritising the mental health of children and young people throughout the COVID-19 pandemic and beyond is critical. These findings highlight that there is wide variation in how children and young people have been affected, with some finding life easier but others experiencing more difficulties. Our findings have identified some sources of variation but we need to continue to gain a better understanding of which families are struggling and what they need to help direct the right advice and support going forward to ensure that this does not have long-lasting consequences.'

The Co-SPACE (COVID-19 Supporting Parents, Adolescents, and Children in Epidemics) survey is still open and keen for parents and carers to share their experiences at: www.cospaceoxford.com/survey.

This research is tracking children and young people's mental health throughout the COVID-19 crisis. Survey results are helping researchers identify what protects children and young people from deteriorating mental health, over time, and at particular stress points, and how this may vary according to child and family characteristics. This will help to identify what advice, support and help parents would find most useful.

Provided by University of Oxford

RE: Mental health issues in children amidst COVID-19 pandemic

Siddharth Madan, Assistant Professor, Department of Ophthalmology, Lady Hardinge Medical college and Associated Hospitals, Department of Ophthalmology, Lady Hardinge Medical college and Associated Hospitals, University of Delhi, New Delhi, India

27 June 2020

Novel coronavirus disease (COVID-19) has affected society in many ways.1 It has brought about lockdown, and school closures have impacted over 1.5 billion children. Restriction in movement, loss of daily wages and isolation invites high levels of stress and anxiety. The repercussion of this is an increased level of psychological and sexual abuse on children at home apart from physical violence. This is more common in parts of society who face the financial crunch as a result of this evolving pandemic. About 30% of children or their parents who are subjected to quarantine at home or are isolated, suffer from acute stress disorder, depression and adjustment problems caused as result of such dramatic change in their daily lifestyle.2 A few college students have experienced suicidal tendencies as well. Lack of companionship and separation from caregivers affects the psychological wellbeing of children. These stressors may trigger new symptoms or exacerbate underlying mental or neurological conditions and may cause sleeping difficulties.3 Although online platforms have been popularized at the right time to prevent academic delays and allow the learning activities of the children to continue, they come with potential risks of unnecessary cyber usage and involvement in irrational behavioral activities. World Health Organization has suggested the importance of reassurance and counseling the younger population, keeping them informed of the facts of the pandemic in a way that is appropriate for their age and as to what is going on now and how they can be protected and disease transmission reduced.4 There is definitely a significant risk of disruption to routine immunization activities due to the COVID-19 related burden on the health care system. This would manifest as an increase in the likelihood of vaccine preventable diseases (VPDs) like measles and would further impact the health care system. Therefore this aspect must not be neglected by the health authorities amidst this pandemic. Psychosocial crisis prevention and intervention model development by the government and healthcare professionals is needed to avert a pandemic of mental ill health erupting simultaneously.5

Competing Interests: None declared.

References

- . RE: Mental health issues in children amidst COVID-19 pandemic. 2020;:-.
- Wang G, Zhang Y, Zhao J, Zhang J, Jiang F. Mitigate the effects of home confinement on children during the

- World Health Organization. Guiding principles for immunization activities during the COVID-19 pandemic: interim guidance, 26 March 2020. 2020 [cited 2020 Jun 27]; Available from: https://apps.who.int/iris/handle/10665/331590
- Dubey S, Biswas P, Ghosh R, Chatterjee S, Dubey MJ, Chatterjee S, et al. Psychosocial impact of COVID-19. Diabetes Metab Syndr. 2020 May 27;14(5):779–88.

See article »

Int Arch Occup Environ Health (2005) 78: 501–509
DOI 10.1007/s00420-004-0584-4

SHORT COMMUNICATION

Y. Li · H. Tokura · Y.P. Guo · A.S.W. Wong · T. Wong
J. Chung · E. Newton

Effects of wearing N95 and surgical facemasks on heart rate, thermal stress and subjective sensations

Received: 7 June 2004 / Accepted: 6 October 2004 / Published online: 26 May 2005

Abstract ***Aim***: The study was aimed at investigating the effects of wearing N95 and surgical facemasks with and without nano-functional treatments on thermophysiological responses and the subjective perception of discomfort. ***Method***: Five healthy male and five healthy female participants performed intermittent exercise on a treadmill while wearing the protective facemasks in a climate chamber controlled at an air temperature of 25°C and a relative humidity of 70%. Four types of facemasks, including N95 (3M 8210) and surgical facemasks, which were treated with nano-functional materials, were used in the study. ***Results***: (1) The subjects had significantly lower average heart rates when wearing nano-treated and untreated surgical facemasks than when wearing nano-treated and untreated N95 facemasks. (2) The outer surface temperature of both surgical facemasks was significantly higher than that of both N95 facemasks. On the other hand, the microclimate and skin temperatures inside the facemask were significantly lower than those in both N95 facemasks. (3) Both surgical facemasks had significantly higher absolute humidity outside the surface than both N95 facemasks. The absolute humidity inside the surgical facemask was significantly lower than that inside both N95 facemasks. (4) Both surgical facemasks were rated significantly lower for perception of humidity, heat, breath resistance and overall discomfort than both N95 facemasks. The ratings for other sensations, including feeling unfit, tight, itchy, fatigued, odorous and salty, that were obtained while the subjects were wearing the surgical facemasks were significantly lower than when the subjects were wearing the N95 facemasks. (5) Subjective preference for the nano-treated surgical facemasks was the highest. There was significant differences in preference between the nano-treated and untreated surgical facemasks and between the surgical and N95 facemasks. ***Discussion***: We discuss how N95 and surgical facemasks induce significantly different temperature and humidity in the microclimates of the facemasks, which have profound influences on heart rate and thermal stress and subjective perception of discomfort.

Keywords Facemasks · Nano-functional materials · Microclimate inside the facemasks · Subjective perception

Y. Li (✉) · H. Tokura · Y.P. Guo · A.S.W. Wong · E. Newton
Institute of Textiles and Clothing,
The Hong Kong Polytechnic University,
Hung Hom, Kowloon,
Hong Kong
E-mail: tcliyi@polyu.edu.hk
Tel.: +852-2-7666479
Fax: +852-2-7731432

T. Wong · J. Chung
School of Nursing,
The Hong Kong Polytechnic University, Kowloon,
Hong Kong

Introduction

Facemasks are critical components of personal protective equipment (PPE) for healthcare workers, particularly when those workers are dealing with transmitted diseases, such as the severe acute respiratory syndrome (SARS) outbreak that occurred in March 2003. Seto et al. (2003) performed a case study in five Hong Kong hospitals, involving 241 non-infected staff and 13 infected staff who were exposed to 11 patients with SARS, and they concluded that SARS was contagious by droplets. They suggested that the wearing of facemasks was of significance in reducing the risk of contagion after exposure to patients with SARS. Wong et al. (2004) reported a study on effective personal protective clothing (PPC) for healthcare workers attending patients with SARS. In the World Health Organization (WHO) (2003) and the US Centers for Disease Control (CDC) (2004) guidelines for PPE, facemasks with 95% filtration efficiency or above are required for healthcare workers exposed to SARS patients.

Hayashi and Tokura (2004) found that it was important to prevent an excessive increase of microclimate temperature and humidity inside the facemask in order to reduce heat stress on the body when farmers were spraying pesticides in a warm environment. Farquharson and Baguley (2003) reported that Emergency Department (ED) staff taking care of SARS patients at a hospital in Toronto wore double isolation gowns, a hair cap, an N95 facemask, a face shield and two pairs of gloves. ED staff had 12-h shift work while wearing N95 facemasks. Only one individual could take off his or her facemask at one time in an enclosed room. As soon as the staff had finished meals and drinks they had to wear the facemask again. Such situations made ED staff extremely stressed. Nielsen et al. (1987) found that the facemask air temperature significantly influenced thermal sensations of the whole body. Meyer et al. (1997) reported that the acceptable duration of wearing respiratory protective devices was about 1 h in a work environment with an air temperature of 18°C on average, and that the comfort sensation was reduced with increase of the air temperature. Similarly, White et al. (1991) found that the wearing of chemical protective clothing significantly reduced acceptable working time due to increased heat stress. These findings show clearly that serious heat stress occurs within the body when protective clothing is worn, which could cause workers to tire more easily and reduce their working time.

In a previous paper we reported that the N95 facemask had a filtration efficiency greater than 96% during wear, comparing surgical facemasks of 95% filtration efficiency (Li et al. 2004). Both N95 and surgical facemasks treated with nano-functional materials had significantly higher repellence to water, which can prevent droplets contaminated with viruses and bacteria from penetrating the facemasks by capillary actions during breathing cycles. Further, it has been shown that surgical facemasks treated with nano-functional materials have a significant ability to inactivate bacteria (Yao et al. 2004). It is important for one to know what impact the wearing of different types of facemasks has on heat stress and discomfort, as the filtration efficiency is similar between surgical and N95 facemasks, and whether the nano-functional treatment has an influence on heat stress and discomfort. In this paper, we report an experimental study on the effects of wearing different kinds of facemasks with and without nano-functional treatments on thermophysiological response and subjective perception of discomfort.

Methods

Subject

Ten healthy subjects, five men and five women, participated in the study, and their physical characteristics are summarized in Table 1. None was a smoker. Female subjects participated in the experiment only when they were during follicular phases.

Every participant was tested four times at the same time of day on four different days, wearing one of four types of facemasks. Before the first experiment the subjects were required to read an information sheet, on which the nature, purpose, method, and risks of the study were described, and then sign a consent form. They had the right to question any part of the procedure and to withdraw themselves from the experiment at any time without penalty. The human subjects ethics and sub-committee of The Hong Kong Polytechnic University approved the experimental protocol.

Facemasks

In the experiments we used four types of facemasks, including N95 (3M 8210) and surgical facemasks, which were treated with nano-functional materials to stop virus penetration by capillary action and to inactivate bacteria (Yao et al. 2004). Both facemasks are commercially available to hospitals and clinics in Hong Kong. The physical characteristics of the four types of mask are described in Table 2.

Physiological measurements

Skin and clothing microclimate (temperature, humidity) inside and outside the facemasks and inside shirts were continuously recorded by a logger (SCXI-1161, National

Table 1 Physical characteristic of human subjects

Characteristic	Male			Female		
	Average	SD	Range	Average	SD	Range
Age (years)	28.0	5.4	22–37	29.4	8.4	21–41
Weight (kg)	68.8	7.8	56–74	55.5	8.9	41–62
Height (cm)	172.5	6.8	164–180	168.2	7.4	151–170

Table 2 Physical characteristics of the masks

Mask type	Treatment	Size (cm)	Materials	Weight (g)	Thickness (mm)
N95	Untreated[a]	φ 12.5×13.2	Coverings: polypropylene	8.99	3.87
N95	Nano-treated[b]		and polyester	9.64	5.17
			Filter media: polypropylene		
Surgical	Untreated[a]	17.3×15.8	Outer and inner layers:	3.26	0.80
Surgical	Nano-treated[b]		polypropylene	3.39	0.85
			Middle layer: melt-blown		

[a]Normal facemasks
[b]Facemasks treated with nano-functional materials

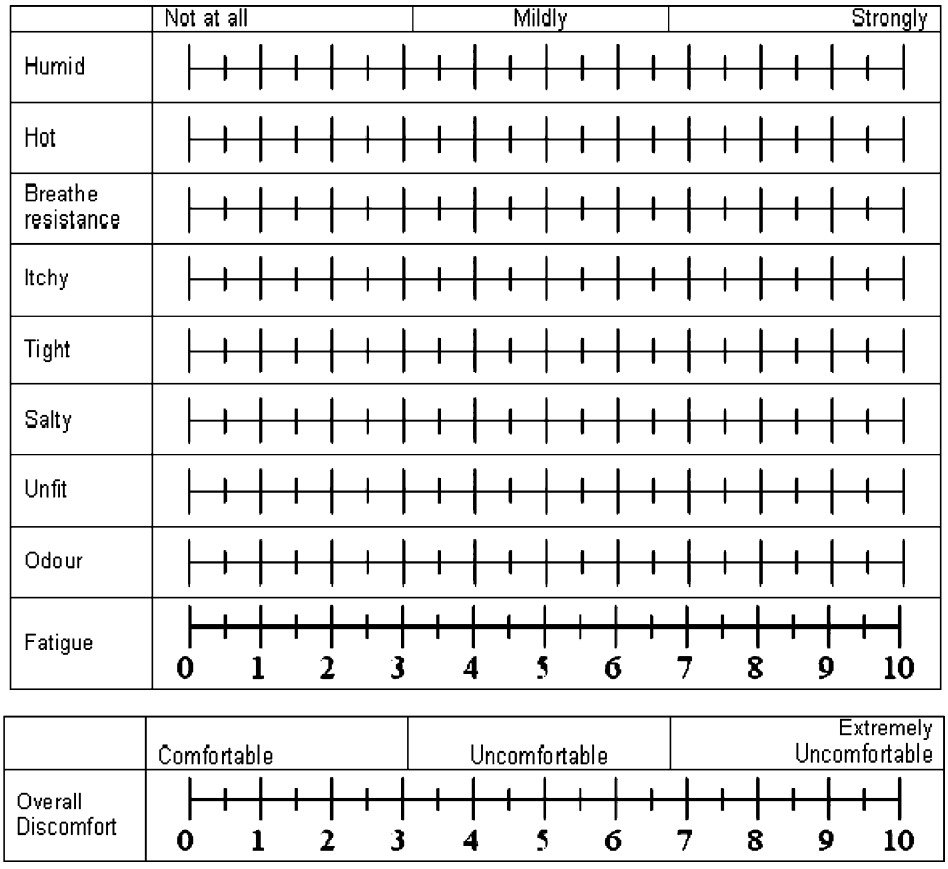


Table 3 Scale of measuring subjective perceptions

Instruments, USA) every 30 s. Sensors for the measurements of temperature and humidity inside shirts were fixed on the left and right chest regions. One uncovered sensor was attached directly to the skin. Facemask microclimate (temperature, humidity) and cheek skin temperature inside the facemasks were measured at the right cheek. Facemask microclimate (temperature, humidity) outside the facemasks was also measured at the right cheek. At the end of each exercise and rest period, heart rate and blood pressure were

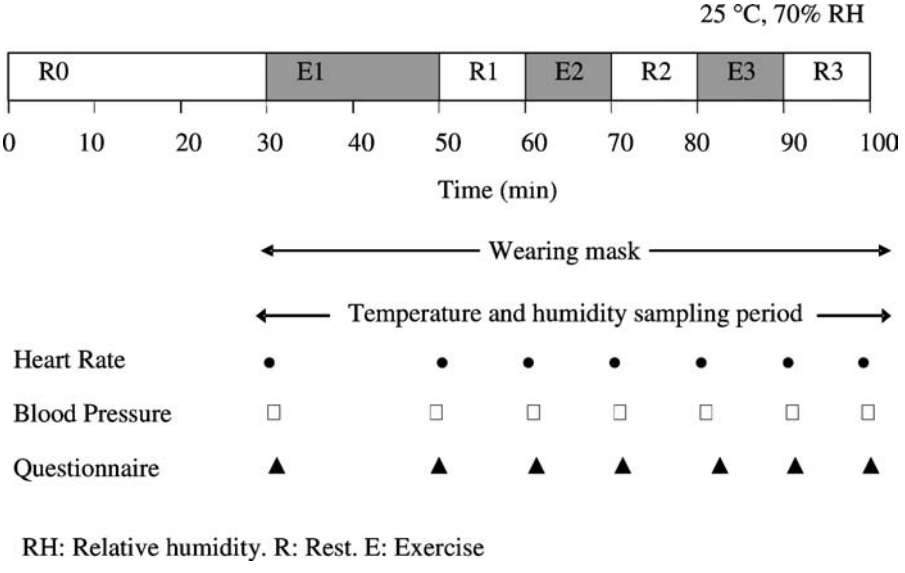


Table 4 The experiment schedule

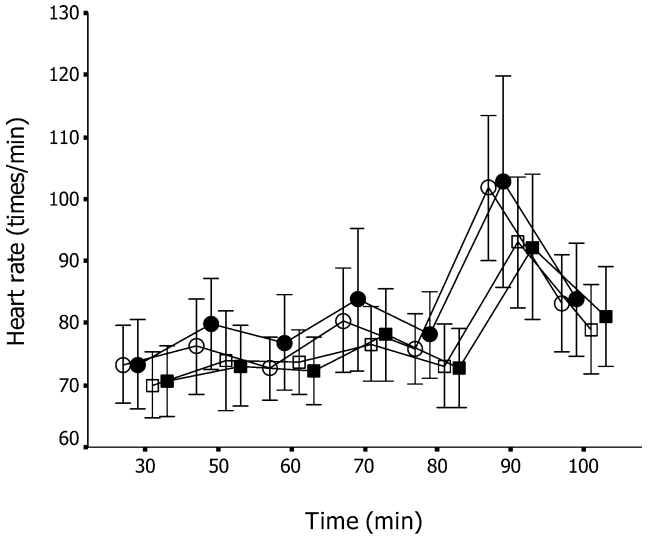


Fig. 1 Temporal changes in mean heart rate under the influence of the four kinds of facemasks. *Open circles* N95 facemask; *closed circles* nano-treated N95 facemask; *open squares* surgical facemask; *closed squares* nano-treated surgical facemask

measured with an upper-arm blood pressure meter (EW 3100, BMEW Ltd., Beijing).

Perception of discomfort

Subjects were required to rate their perceptions of ten sensations of discomfort: humidity, heat, breathing resistance, itchiness, tightness, saltiness, feeling unfit, odor, fatigue, and overall discomfort, at 30, 50, 60, 70, 80, 90 and 100 min. Table 3 shows the rating scales used by the subjects. In addition, at 100 min, the subjects were asked to reply to the question "How do you like the facemask?" by rating on a scale ranging from 0 to 10, with 0 representing "not at all", 5 representing "acceptable" and 10 representing "very fond of". This rating was used to obtain the preference of subjects for the four kinds of facemasks.

Experimental protocol

The experiments were carried out for 3 months from May to July. They were performed twice a day, one from 0900 h to 1100 h and another from 1500 h to 1700 h. The experimental protocol was randomized for men and women, and for the four types of facemasks.

A subject entered the climate chamber controlled at an air temperature of 25°C and a relative humidity of 70%, which is similar to the conditions in the hospitals. After body mass had been measured, the subject wore a 100% cotton T-shirt, short pants and sports sandals. Sensors were attached to different areas with surgical tape. Following a rest for 30 min on a chair (R0), during which time the subject was required to drink 500 ml water, the subject voided the bladder completely and put

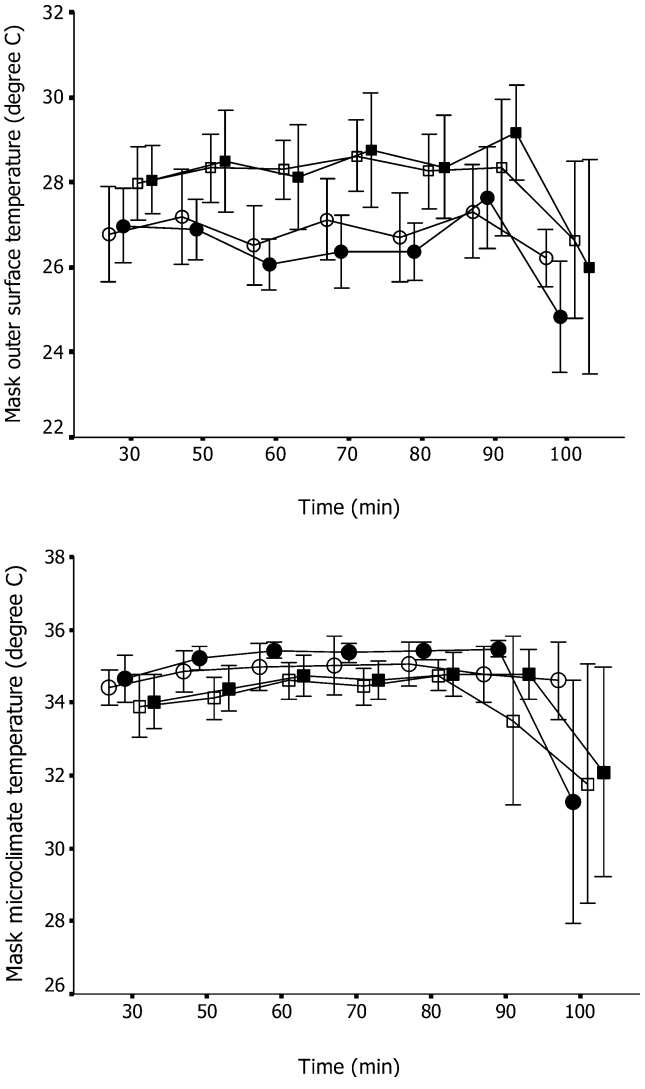


Fig. 2 Temporal change in mean temperature on the outer surface of the facemasks (*top*) and in the microclimate of the facemasks (*bottom*) under the influence of the four kinds of facemasks. *Open circles* N95 facemask; *closed circles* nano-treated N95 facemask; *open squares* surgical facemask; *closed squares* nano-treated surgical facemask

on a facemask, randomly selected. Then, the subject walked for 20 min at 3.2 km/h (E1) and took a rest for 10 min (R1); walked for 10 min at 4.8 km/h (E2) and took a rest for 10 min again (R2); and finally, the subject walked for 10 min at 6.4 km/h (E3) and took a rest for 10 min (R3). These workloads resembled approximately those performed by healthcare workers in a hospital ward. The schedule of the experiment is shown in Table 4. The participant took off the mask at 100 min, completing the whole experiment.

Statistical analysis

As mask microclimate temperature is a key parameter indicating thermal stress, we used this parameter to estimate the sample size. According to previous reports,

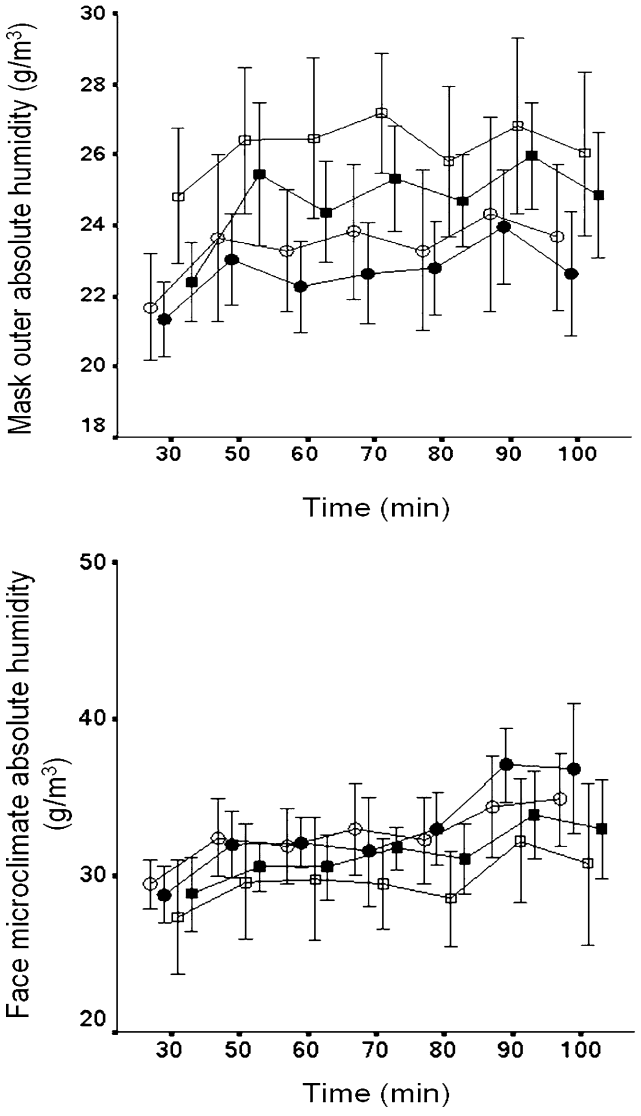


Fig. 3 Temporal changes in mean absolute humidity outside (*top*) and inside (*bottom*) the four kinds of facemasks. *Open circles* N95 facemask; *closed circles* nano-treated N95 facemask; *open squares* surgical facemask; *closed squares* nano-treated surgical facemask

the difference in microclimate temperature between masks is approximately 0.9°C and standard deviation is around 0.5°C (Hayashi and Tokura 2004). From this assumption, a sample size calculation reveals that ten participants are enough to reach an error of probability of $< 5\%$ and a power of 90%.

Physiological parameters (including heart rate, temperature and humidity) and psychological responses (including perception of humidity, heat and breath resistance) were analyzed statistically. The influence of time, facemask type, nano-treatment, and their interactions on these human physiological and psychological responses were investigated by analysis of variance (ANOVA) and *t*-tests to determine whether the above factors had significant effect on the measured parameters.

Results

Physiological parameters

Heart rate

Figure 1 compares temporal changes of mean heart rates when the subjects were wearing the four kinds of facemasks. The pattern of changes in mean heart rate amongst these facemasks is similar, reaching peaks at the end of the third exercise session. The subjects had lower mean heart rates when wearing nano-treated and untreated surgical masks than when wearing nano-treated and untreated N95 facemasks. Significant differences were found among the four kinds of facemasks at the level of $P < 0.01$ ($F = 10.76$).

Temperature and humidity

Mask microclimate and face skin temperatures

Figure 2 shows temporal changes in temperatures on the facemasks' outer surfaces and in the facemasks' microclimates. The outer surface temperatures of both surgical facemasks were significantly higher than those of both N95 facemasks ($F = 94.4$, $P < 0.01$) (top of Fig. 3). On the other hand, microclimate temperatures inside the mask were significantly lower in both surgical masks than those in both N95 facemasks ($F = 25.7$, $P < 0.01$) (bottom of Fig. 3). The skin temperatures inside both surgical facemasks were significantly lower than those in both N95 facemasks ($F = 40.7$, $P < 0.01$).

Humidity outside and inside the facemask

Figure 3 (top) shows that both surgical facemasks had significantly higher absolute humidity on the outside surface than both N95 facemasks ($F = 6.9$, $P < 0.01$). The overall mean absolute humidity ± SD in nano-treated and untreated surgical facemasks was 24.7 ± 2.76 g/m^3 and 26.2 ± 2.74 g/m^3, respectively. The overall mean absolute humidity ± SD in nano-treated and untreated N95 facemasks was 22.7 ± 1.83 g/m^3 and 23.4 ± 2.74 g/m^3, respectively. Figure 3 (bottom) shows that the absolute microclimate humidity inside the surgical mask was significantly lower than inside both N95 facemasks. The overall mean absolute humidity ± SD in nano-treated and untreated surgical facemasks was 30.2 ± 4.32 g/m^3 and 28.64 ± 5.37 g/m^3, respectively. The overall mean absolute humidity ± SD in nano-treated and untreated N95 facemasks was 31.2 ± 5.47 g/m^3 and 31.8 ± 4.17 g/m^3, respectively.

Table 5 summarizes the influences of time, facemask, nano-treatment, and their interactions on physiological parameters (heart rate, blood pressure) and microclimate (temperature, absolute humidity) by ANOVA. For each parameter a multi-way analysis of variances was

Table 5 Influences of time, facemask, nano-treatment, and their interactions on physiological parameters. $P > 0.05$ is considered as being not significant and is shown as a dash. *Mask* type of facemask, *Treat* nano-treatment

Physiological parameters	*P* values						
	Time	Mask	Treat	Time × Mask	Time × Treat	Mask × Treat	Time × Mask × Treat
Heart rate	0.000	0.001	–	–	–	0.000	–
Diastolic blood pressure	–	–	–	–	–	–	–
Systolic blood pressure	0.000	–	–	–	–	–	–
Mask outer humidity	0.000	0.000	0.000	–	–	–	–
Face microclimate humidity	0.000	0.000	0.035	–	–	–	–
Chest microclimate humidity	0.000	–	–	–	–	0.009	–
Mask outside temperature	0.030	0.000	–	–	–	–	–
Face microclimate temperature	0.000	0.000	0.003	–	–	–	–
Face skin temperature	0.002	0.000	0.000	–	–	0.005	–
Chest microclimate temperature	–	–	–	–	–	0.039	–

carried out to identify the statistical significance of the influences of the three variables: time, type of facemasks and nano-treatment, as well as their interactions. To save space, only the *P* values are used to show the statistical significance. A $P > 0.05$ is considered as being not significant and is shown as a dash, and a $P < 0.0005$ is considered as being significant and is shown as "0.000". Nine parameters, including heart rate, systolic blood pressure, absolute humidity (mask outer surface, face microclimate, left chest microclimate and right chest skin) and temperature (mask outer surface, face microclimate and face skin) were significantly influenced by time. Other factors that had significant effect on the measured parameters were mask, interaction of mask and nano-treatment and nano-treatment on its own.

Subjective ratings

Figure 4 compares subjective ratings for thermal sensation and overall discomfort for the four types of facemasks. In general, the ratings for humidity, heat, breath

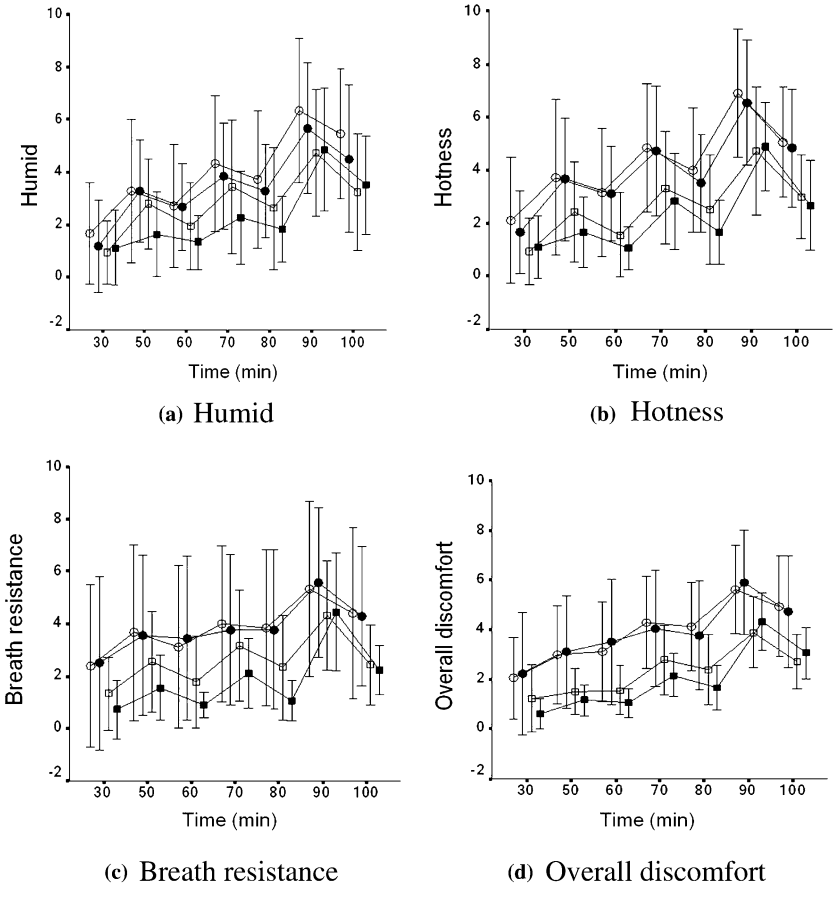


Fig. 4 Subjective ratings for various sensations under the influence of the four kinds of facemasks: **a** humidity, **b** heat, **c** breath resistance, **d** overall discomfort. *Open circles* N95 facemask; *closed circles* nano-treated N95 facemask; *open squares* surgical facemask; *closed squares* nano-treated surgical facemask

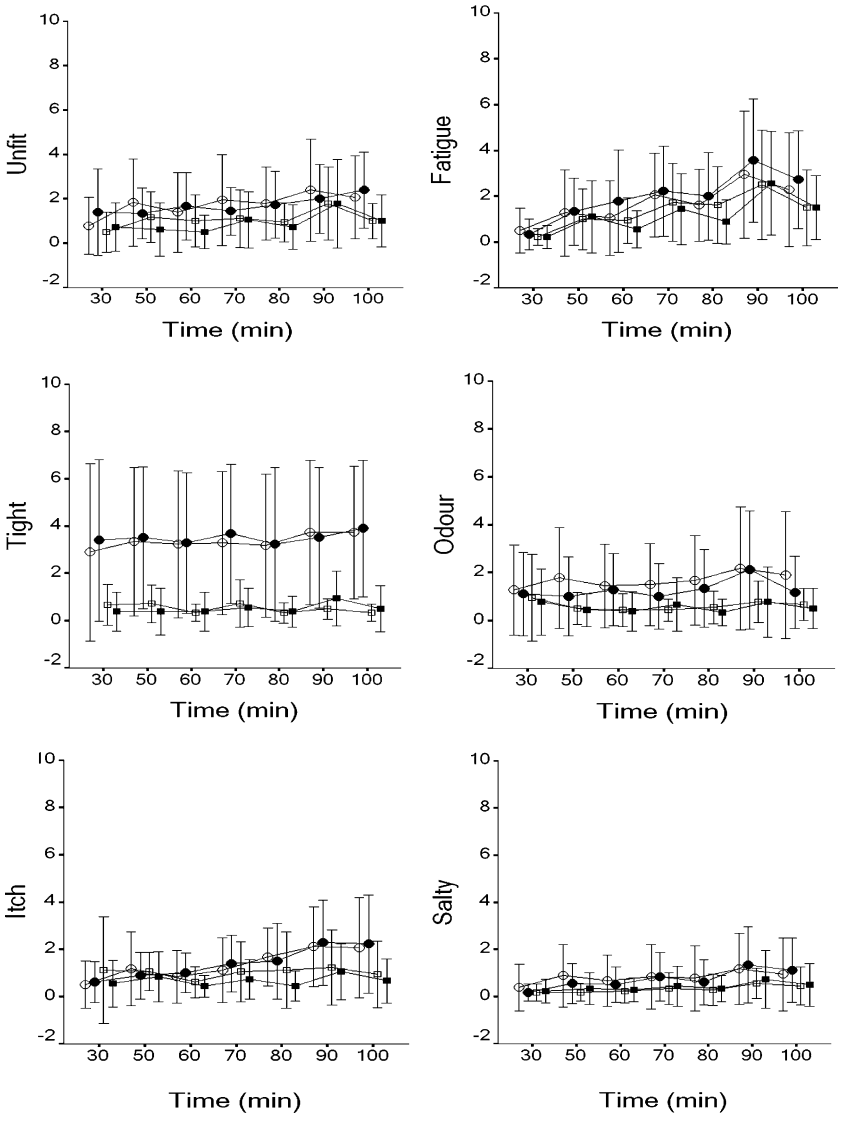


Fig. 5 Others sensations under the influence of the four kinds of facemasks: feeling unfit, tight, itchy, fatigued, odorous and salty. *Open circles* N95 facemask; *closed circles* nano-treated N95 facemask; *open squares* surgical facemask; *closed squares* nano-treated surgical facemask

resistance and overall discomfort increased gradually with time and increase of workload. Facemask type had great influence on the perception of humidity ($F = 6.9$, $P < 0.01$), heat ($F = 15.4$, $P < 0.01$), breath resistance ($F = 15.0$, $P < 0.01$) and overall discomfort ($F = 23.1$, $P < 0.01$). Both surgical facemasks had significantly lower ratings than the two N95 facemasks, which suggested that when wearing either of the surgical

Table 6 Influences of time, facemask, nano-treatment, and their interactions on various subjective sensations. $P > 0.05$ is considered as being not significant and is shown as a dash. *Mask* type of facemask, *Treat* nano-treatment

Subjective sensations	*P* values						
	Time	Mask	Treat	Time × Mask	Time × Treat	Mask × Treat	Time × Mask × Treat
Humidity	0.000	0.000	–	–	–	–	–
Heat	0.000	0.000	–	–	–	–	–
Breath resistance	0.000	0.000	–	–	–	–	–
Itchy	0.017	0.002	–	–	–	–	–
Tight	–	0.000	–	–	–	–	–
Salty	–	0.001	–	–	–	–	–
Feeling unfit	0.047	0.000	–	–	–	–	–
Odorous	–	0.000	–	–	–	–	–
Fatiguing	0.000	0.011	–	–	–	–	–
Overall discomfort	0.000	0.000	–	–	–	–	–

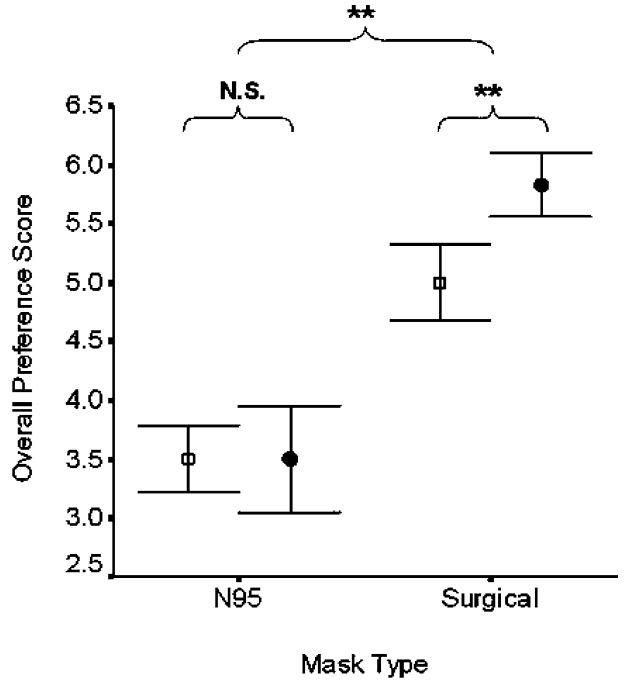


Fig. 6 Subjective preferences for the four kinds of facemasks

facemasks the subject felt drier, cooler, more able to breathe easily and less uncomfortable than when wearing either of the N95 facemasks. The ratings for humidity, heat, breathing resistance and discomfort of facemasks treated with nano-functional materials appear lower than those for untreated facemasks but are not statistically significant.

Figure 5 shows the subjective ratings for other sensations obtained while the subjects were wearing the facemasks. There are significant differences in the subjective perceptions feeling unfit ($F = 5.3$, $P < 0.01$), tight ($F = 34.6$, $P < 0.01$), itchy ($F = 4.7$, $P < 0.01$), fatigued ($F = 2.7$, $P < 0.05$), odorous ($F = 7.9$, $P < 0.01$) and salty ($F = 3.9$, $P < 0.01$). The ratings for those sensations were significantly lower when the subjects were wearing the surgical facemasks than when they were wearing either of the N95 facemasks, showing that the subjects felt less unfit, less tight, less itchy, less fatigued, less odorous and less salty with the surgical facemasks than with the N95 masks.

Table 6 summarizes the result of ANOVA, which show the influences of time, facemask, nano-treatment, and their interactions on subjective ratings for individual sensations and overall discomfort. Again, for each sensation, we carried out a multi-way analysis of variances to identify the statistical significance of the influences of the three variables: time, type of facemasks and nano-treatment, as well as their interactions. To save space, only the P values are used to show the statistical significance. A $P > 0.05$ is considered as being not significant and is shown as a dash, and a $P < 0.0005$ is considered as being significant and is marked as "0.000". As shown in Table 6, facemask type influences subjects' perception of all the nine individual sensations and overall discomfort significantly ($P < 0.05$). On the other hand, all sensations were not significantly influenced by time and nano-treatment. There were no significant differences between ratings for tight, salty and odorous at different time periods.

Figure 6 shows the preferences of subjects for the four kinds of facemasks. Subjective preference for the nano-treated surgical facemasks is the highest, followed by the untreated surgical masks, the nano-treated N95 and then the untreated N95 facemask. There is a significant difference in preference between the nano-treated and untreated surgical facemasks and between the surgical and N95 facemasks. There is no significant difference in subjective preference between nano-treated and untreated N95 facemasks.

Discussion and conclusion

The results from the experiment demonstrate that heart rate, microclimate (temperature, humidity) and subjective ratings were significantly influenced by the wearing of different kinds of facemasks. Nielsen et al. (1987) observed that delivery of air with different temperatures into a facemask corresponded to the application of a local thermal stimulus to the skin surface around the mouth, nose and cheek. This local thermal stimulus also affected the heat exchange from the respiratory tract. In our investigation, microclimate temperature, humidity and skin temperature inside the facemask increased with the start of step exercise, which led to the different perceptions of humidity, heat and high breathing resistance among the subjects wearing the facemasks. High breathing resistance made it difficult for the subject to breathe and take in sufficient oxygen. Shortage of oxygen stimulates the sympathetic nervous system and increases heart rate (Ganong 1997). It was probable that the subjects felt unfit, fatigued and overall discomfort due to this reason. White et al. (1991) found that the increases in heart rate, skin temperature and subjective ratings may pose substantial additional stress to the wearer and might reduce work tolerance. This could be the reason why Farquharson reported that working 12-h shifts while wearing an N95 mask had indeed been a challenge to their ED staff (Farquharson and Baguley 2003).

Significant differences were observed between N95 and surgical masks. Mean heart rate, microclimate temperature, humidity and skin temperature inside the facemask, together with perceived humidity, heat, breathing resistance in the facemask, and itchiness, fatigue and overall discomfort, were significantly ($P < 0.01$) higher for N95 masks than for surgical masks. In other words, the subjective perception of breathing difficulty and discomfort increased significantly with increasing thermal stress. This finding agrees with the observations reported by White et al. (1991). The surface

temperature outside the facemask was lower, and the temperature in the facemask microclimate was significantly higher, for the N95 masks than for the surgical masks (Fig. 3), indicating that the heat loss from the respiratory tract is more difficult to endure in N95 masks, inducing higher heat stress and perception of discomfort. This agrees well with the observations reported by Hayashi and Tokura (2004).

As the purpose of wearing the facemasks is to protect the wearers by filtering out viruses and bacteria, it is obviously questionable whether the surgical masks, which induce less heat stress and discomfort, can provide enough protection for healthcare workers. As reported previously, the in vivo filtration efficiency and physical properties of the masks were investigated at the same time (Li et al., unpublished data). During the simulation wear trials, in vivo filtration efficiency of N95 facemasks was 96%, in comparison with 95% for surgical facemasks. Furthermore, the surgical facemasks with significantly higher moisture permeability and air-permeability were thinner than the N95 facemasks, indicating that surgical facemasks should be more breathable and less humid and hot, which agrees with the in vivo measurements of temperature and humidity inside and outside the masks and the subjects' perception of breathing resistance and discomfort.

It is interesting to note that no significant difference was found between nano-treated and untreated facemasks for physiological measurements and subjective perceptions, even though nano-treated surgical and N95 facemasks were perceived to be slightly less uncomfortable. On the other hand, subjective preferences for the nano-treated surgical masks were significantly higher than those for the untreated surgical facemasks. This indicates that the nano-functional treatment of surgical and N95 facemasks does not have significant negative effects on the thermophysiological responses and subjective perceptions of discomfort.

Therefore, it can be concluded that N95 and surgical facemasks can induce significantly different temperatures and humidity in the microclimates of facemasks, which have profound influences on heart rate and thermal stress and subjective perception of discomfort.

Acknowledgments We would like to express our thanks to the Hong Kong Polytechnic University for supporting this research through projects A188, YD56 and G-YD80.

References

Farquharson C, Baguley K (2003) Responding to the severe acute respiratory syndrome (SARS) outbreak: lessons learned in a Toronto emergency department. J Emerg Nurs 23:222–228

Ganong WF (1997) Review of Medical Physiology. Appleton and Lange, Stamford, pp 565–566

Hayashi C, Tokura H (2004) The effects of two kinds of masks (with/without exhaust valve) on clothing microclimates inside the mask in participants wearing protective clothing for spraying pesticides. Int Arch Occup Environ Health 77:73–78

Li Y, Chung JWY, Wong TKS, Newton N, Hu JY, Guang YT, Guo YP, Yao L, Song QW (2004) In-vivo protective performance of facemasks coated with nano-functional materials. Hong Kong SARS Forum and Hospital Authority Convention, 8–11 May, Hong Kong, 118 pp

Meyer JP, Héry M, Herrault J, Hubert G, François D, Hecht G, Villa M (1997) Field study of subjective assessment of negative pressure half-masks. Influence of the work conditions on comfort and efficiency. Appl Ergon 28:331–338

Nielsen R, Berglund LG, Gwosdow AR, DuBois AB (1987) Thermal sensation of the body as influenced by the thermal microclimate in a face mask. Ergonomics 30:1689–1703

Seto WH, Tsang D, Yung RWH, Ching TY, Ng TK, Ho M, Ho LM (2003) Effectiveness of precautions against droplets and contact in prevention of nosocomial transmission of severe acute respiratory syndrome (SARS). Lancet 361:1519–1520

US Centers for Disease Control (CDC) (2004) Infection control in healthcare, home, and community settings, 8 January

White MK, Hodous TK, Vercruyssen M (1991) Effects of thermal environment and chemical protective clothing on work tolerance, physiological responses, and subjective ratings. Ergonomics 34:445–457

WHO (2003) Hospital infection control guidance for severe acute respiratory syndrome (SARS). World Health Organization, Geneva

Wong TKS, Chung JWY, Li Y, Chan WF, Ching PTY, Lam CHS, Chow CB, Seto WH (2004) Effective personal protective clothing (PPC) for healthcare workers attending patients with severe acute respiratory syndrome (SARS). Am J Infect Control 32:90–96

Yao L, Li Y, Leung P, Song QW (2004) In-vivo antibacterial activity of surgical facemasks coated with nano-functional materials. Hong Kong SARS Forum and Hospital Authority Convention, 8–11 May, Hong Kong, p 119

ScienceDaily®

Your source for the latest research news

Immune study points to new ways to treat lung disease

Date: August 14, 2017

Source: University of Edinburgh

Summary: Fresh insight into how the immune system keeps itself in check could lead to new ways of fighting chronic lung disease, report investigators.

FULL STORY

Fresh insight into how the immune system keeps itself in check could lead to new ways of fighting chronic lung disease.

New findings could open avenues of research for tackling damage caused by cells that overreact to infection.

Scientists from the University of Edinburgh studied immune cells known as neutrophils, which fight bacteria and help to cause inflammation, a normal biological response to wounds or infection that is recognisable by redness and swelling.

They found that when neutrophils lose a certain oxygen-sensing protein, the cells become overactive and respond excessively to infection in a harmful way.

Studies in mice found that by preventing the cells from using sugar, this effect could be reversed.

Studying the effects of oxygen-sensing proteins in immune cells is especially relevant for patients who often have low levels of oxygen in their body and chronic lung inflammation.

The study, funded by the Wellcome Trust, is published in the *Journal of Clinical Immunology*.

Professor Sarah Walmsley, of the MRC Centre for Inflammation Research, said: "This finding demonstrates the therapeutic potential of targeting how neutrophils use glucose in the treatment of chronic inflammatory diseases. As many of these diseases have no effective treatment, future studies examining the role of glucose in regulating neutrophils and inflammation are critical."

Story Source:

Materials provided by **University of Edinburgh**. *Note: Content may be edited for style and length.*

Journal Reference:

1. Pranvera Sadiku, Joseph A. Willson, Rebecca S. Dickinson, Fiona Murphy, Alison J. Harris, Amy Lewis, David Sammut, Ananda S. Mirchandani, Eilise Ryan, Emily R. Watts, A.A. Roger Thompson, Helen M. Marriott, David H. Dockrell, Cormac T. Taylor, Martin Schneider, Patrick H. Maxwell, Edwin R. Chilvers, Massimilliano Mazzone, Veronica Moral, Chris W. Pugh, Peter J. Ratcliffe, Christopher J. Schofield, Bart Ghesquiere, Peter

Carmeliet, Moira K.B. Whyte, Sarah R. Walmsley. **Prolyl hydroxylase 2 inactivation enhances glycogen storage and promotes excessive neutrophilic responses**. *Journal of Clinical Investigation*, 2017; DOI: 10.1172/JCI90848

RELATED STORIES

Human Immune Response in the Fruit Fly

June 19, 2018 — Researchers have seen how both humans and fruit flies deploy a protein that a plays a critical role in their immune responses to invading bacteria. The discovery gives scientists evolutionary insight ...

The Swollen Colon: Cause of Chronic Inflammation Discovered

Apr. 28, 2017 — Too much of the oncogene Bcl-3 leads to chronic intestinal diseases, report investigators. They describe in a new report exactly how it throws the immune system ...

E-Cigarettes Popular Among Smokers With Existing Illnesses

Feb. 21, 2017 — Current and former smokers suffering from illnesses like chronic lung or cardiovascular disease are more likely to use e-cigarettes, investigators ...

New Technique Uses Immune Cells to Deliver Anti-Cancer Drugs

Jan. 4, 2017 — Some researchers are working to discover new, safer ways to deliver cancer-fighting drugs to tumors without damaging healthy cells. Others are finding ways to boost the body's own immune system to ...

Published in final edited form as:
Psychol Bull. 2004 July ; 130(4): 601–630.

Psychological Stress and the Human Immune System: A Meta-Analytic Study of 30 Years of Inquiry

Suzanne C. Segerstrom and
University of Kentucky

Gregory E. Miller
University of British Columbia

Abstract

The present report meta-analyzes more than 300 empirical articles describing a relationship between psychological stress and parameters of the immune system in human participants. Acute stressors (lasting minutes) were associated with potentially adaptive upregulation of some parameters of natural immunity and downregulation of some functions of specific immunity. Brief naturalistic stressors (such as exams) tended to suppress cellular immunity while preserving humoral immunity. Chronic stressors were associated with suppression of both cellular and humoral measures. Effects of event sequences varied according to the kind of event (trauma vs. loss). Subjective reports of stress generally did not associate with immune change. In some cases, physical vulnerability as a function of age or disease also increased vulnerability to immune change during stressors.

Since the dawn of time, organisms have been subject to evolutionary pressure from the environment. The ability to respond to environmental threats or stressors such as predation or natural disaster enhanced survival and therefore reproductive capacity, and physiological responses that supported such responses could be selected for. In mammals, these responses include changes that increase the delivery of oxygen and glucose to the heart and the large skeletal muscles. The result is physiological support for adaptive behaviors such as “fight or flight.” Immune responses to stressful situations may be part of these adaptive responses because, in addition to the risk inherent in the situation (e.g., a predator), fighting and fleeing carries the risk of injury and subsequent entry of infectious agents into the bloodstream or skin. Any wound in the skin is likely to contain pathogens that could multiply and cause infection (Williams & Leaper, 1998). Stress-induced changes in the immune system that could accelerate wound repair and help prevent infections from taking hold would therefore be adaptive and selected along with other physiological changes that increased evolutionary fitness.

Modern humans rarely encounter many of the stimuli that commonly evoked fight-or-flight responses for their ancestors, such as predation or inclement weather without protection. However, human physiological response continues to reflect the demands of earlier environments. Threats that do not require a physical response (e.g., academic exams) may therefore have physical consequences, including changes in the immune system. Indeed, over the past 30 years, more than 300 studies have been done on stress and immunity in humans, and together they have shown that psychological challenges are capable of modifying various features of the immune response. In this article we attempt to consolidate empirical knowledge about psychological stress and the human immune system through meta-analysis. Both the construct of stress and the human immune system are complex, and both could consume book-

Correspondence concerning this article should be addressed to Suzanne C. Segerstrom, Department of Psychology, University of Kentucky, 115 Kastle Hall, Lexington, KY 40506-0044, or Gregory E. Miller, Department of Psychology, University of British Columbia, 2136 West Mall, Vancou-ver, British Columbia V6T IZ4, Canada. E-mail: scsege0@uky.edu or gemiller@psych.ubc.ca.

length reviews. Our review, therefore, focuses on those aspects that are most often represented in the stress and immunity literature and therefore directly relevant to the meta-analysis.

Conceptualizing Stress

Despite nearly a century of research on various aspects of stress, investigators still find it difficult to achieve consensus on a satisfactory definition of this concept. Most of the studies contributing to this review simply define *stress* as circumstances that most people would find stressful, that is, stressors. We adopted Elliot and Eisdorfer's (1982) taxonomy to characterize these stressors. This taxonomy has the advantage of distinguishing among stressors on two important dimensions: duration and course (e.g., discrete vs. continuous). The taxonomy includes five categories of stressors. *Acute time-limited stressors* involve laboratory challenges such as public speaking or mental arithmetic. *Brief naturalistic stressors*, such as academic examinations, involve a person confronting a real-life short-term challenge. In *stressful event sequences*, a focal event, such as the loss of a spouse or a major natural disaster, gives rise to a series of related challenges. Although affected individuals usually do not know exactly when these challenges will subside, they have a clear sense that at some point in the future they will. *Chronic stressors*, unlike the other demands we have described, usually pervade a person's life, forcing him or her to restructure his or her identity or social roles. Another feature of chronic stressors is their stability—the person either does not know whether or when the challenge will end or can be certain that it will never end. Examples of chronic stressors include suffering a traumatic injury that leads to physical disability, providing care for a spouse with severe dementia, or being a refugee forced out of one's native country by war. *Distant stressors* are traumatic experiences that occurred in the distant past yet have the potential to continue modifying immune system function because of their long-lasting cognitive and emotional sequelae (Baum, Cohen, & Hall, 1993). Examples of distant stressors include having been sexually assaulted as a child, having witnessed the death of a fellow soldier during combat, and having been a prisoner of war.

In addition to the presence of difficult circumstances, investigators also use life-event interviews and life-event checklists to capture the total number of different stressors encountered over a specified time frame. Depending on the instrument, the focus of these assessments can be either major life events (e.g., getting divorced, going bankrupt) or minor daily hassles (e.g., getting a speeding ticket, having to clean up a mess in the house). With the more sophisticated instruments, judges then code stressor severity according to how the average person in similar biographical circumstances would respond (e.g., S. Cohen et al., 1998; Evans et al., 1995).

A smaller number of studies enrolled large populations of adults who were not experiencing any specific difficulty and examined whether their immune responses varied according to their reports of perceived stress, intrusive thoughts, or both. Other studies have examined stressed populations, in which a larger range of subjective responses may be detected. This work grows out of the view that people's biological responses to stressful circumstances are heavily dependent on their appraisals of the situation and cognitive and emotional responses to it (Baum et al., 1993; Frankenhauser, 1975; Tomaka, Blascovich, Kibler, & Ernst, 1997).

Overview of the Immune System

As many behavioral scientists are unfamiliar with the details of the immune system, we provide a brief overview. For a more complete treatment, the reader is directed to the sources for the information presented here (Benjamini, Coico, & Sunshine, 2000; Janeway & Travers, 1997; Rabin, 1999). Critical characteristics of various immune components and assays are also listed in Table 1.

Components of the Immune System

There are several useful ways of dividing elements of the immune response. For the purposes of understanding the relationship of psychosocial stressors to the immune system, it is useful to distinguish between *natural* and *specific* immunity. Natural immunity is an immune response that is characteristic not only of mammals but also lower order organisms such as sponges. Cells involved in natural immunity do not provide defense against any particular pathogen; rather, they are all-purpose cells that can attack a number of different pathogens[1] and do so in a relatively short time frame (minutes to hours) when challenged. The largest group of cells involved in natural immunity is the granulocytes. These cells include the *neutrophil* and the *macrophage*, phagocytic cells that, as their name implies, eat their targets. The generalized response mounted by these cells is *inflammation*, in which neutrophils and macrophages congregate at the site of injury or infection, release toxic substances such as oxygen radicals that damage invaders, and phagocytose both invaders and damaged tissue. Macrophages in particular also release communication molecules, or *cytokines*, that have broad effects on the organism, including fever and inflammation, and also promote wound healing. These proinflammatory cytokines include interleukin(IL)-1, IL-6, and tumor necrosis factor alpha (TNFα). Other granulocytes include the mast cell and the eosinophil, which are involved in parasitic defense and allergy.

Another cell involved in natural immunity is the natural killer cell. Natural killer cells recognize the lack of a self-tissue molecule on the surface of cells (characteristic of many kinds of virally infected and some cancerous cells) and lyse those cells by releasing toxic substances on them. Natural killer cells are thought to be important in limiting the early phases of viral infections, before specific immunity becomes effective, and in attacking self-cells that have become malignant.

Finally, complement is a family of proteins involved in natural immunity. Complement protein bound to microorganisms can up-regulate phagocytosis and inflammation. Complement can also aid in antibody-mediated immunity (discussed below as part of the specific immune response).

Specific immunity is characterized by greater specificity and less speed than the natural immune response. Lymphocytes have receptor sites on their cell surfaces. The receptor on each cell fits with one and only one small molecular shape, or antigen, on a given invader and therefore responds to one and only one kind of invader. When activated, these antigen-specific cells divide to create a population of cells with the same antigen specificity in a process called *clonal proliferation*, or the *proliferative response*. Although this process is efficient in terms of the number of cells that have to be supported on a day-to-day basis, it creates a delay of up to several days before a full defense is mounted, and the body must rely on natural immunity to contain the infection during this time.

There are three types of lymphocytes that mediate specific immunity: T-helper cells, T-cytotoxic cells, and B cells. The main function of T-helper cells is to produce cytokines that direct and amplify the rest of the immune response. T-cytotoxic cells recognize antigen expressed by cells that are infected with viruses or otherwise compromised (e.g., cancer cells) and lyse those cells. B cells produce soluble proteins called *antibody* that can perform a number of functions, including neutralizing bacterial toxins, binding to free virus to prevent its entry into cells, and opsonization, in which a coating of antibody increases the effectiveness of natural

[1]The term *pathogen* is used here to refer to microorganisms that can cause disease. This term is most appropriate in the evolutionary context we proposed in the article's introduction because it focuses on susceptibility to infection. However, the reader should be aware that pathogens are only a subset of *antigens*, that is, all substances that evoke an immune response. Other antigenic substances include, for example, transformed self-cells (i.e., cancer cells), transplanted tissue, and allergens (i.e., antigens that evoke an allergic response).

immunity. There are five kinds of antibody: Immunoglobulin (Ig) A is found in secretions, IgE binds to mast cells and is involved in allergy, IgM is a large molecule that clears antigen from the bloodstream, IgG is a smaller antibody that diffuses into tissue and crosses the placenta, and IgD is of unknown significance but may be produced by immature B cells.

An important immunological development is the recognition that specific immunity in humans is composed of cellular and humoral responses. Cellular immune responses are mounted against intracellular pathogens like viruses and are coordinated by a subset of T-helper lymphocytes called *Th1* cells. In the Th1 response, the T-helper cell produces cytokines, including IL-2 and interferon gamma (IFNγ). These cytokines selectively activate T-cytotoxic cells as well as natural killer cells. Humoral immune responses are mounted against extracellular pathogens such as parasites and bacteria; they are coordinated by a subset of T-helper lymphocytes called *Th2* cells. In the Th2 response, the T-helper cell produces different cytokines, including IL-4 and IL-10, which selectively activate B cells and mast cells to combat extracellular pathogens.

Immune Assays

Immune assays can quantify cells, proteins, or functions. The most basic parameter is a simple count of the number of cells of different subtypes (e.g., neutrophils, macrophages), typically from peripheral blood. It is important to have an adequate number of different types of immune cells in the correct proportions. However, the normal range for these enumerative parameters is quite large, so that "correct" numbers and proportions can cover a wide range, and small changes are unlikely to have any clinical significance in healthy humans.

Protein production—either of antibody or cytokines—can be measured in vitro by stimulating cells and measuring protein in the supernatant or in vivo by measuring protein in peripheral blood. For both antibody and cytokine, higher protein production may represent a more robust immune response that can confer protection against disease. Two exceptions are levels of proinflammatory cytokines (IL-1, IL-6, and TNFα) and antibody against latent virus. Proinflammatory cytokines are increased with systemic inflammation, a risk factor for poorer health resulting from cardiac disease, diabetes mellitus, or osteoporosis (Ershler & Keller, 2000; Luster, 1998; Papanicoloaou, Wilder, Manolagas, & Chrousos, 1998). Antibody production against latent virus occurs when viral replication triggers the immune system to produce antibodies in an effort to contain the infection. Most people become infected with latent viruses such as Epstein-Barr virus during adolescence and remain asymptomatically infected for the rest of their lives. Various processes can activate these latent viruses, however, so that they begin actively replicating. These processes may include a breakdown in cellular immune response (Jenkins & Baum, 1995). Higher antibody against latent viruses, therefore, may indicate poorer immune control over the virus.

Functional assays, which are performed in vitro, measure the ability of cells to perform specific activities. In each case, higher values may represent more effective immune function. Neutrophils' function can be quantified by their ability to migrate in a laboratory assay and their ability to release oxygen radicals. The natural killer cytotoxicity assay measures the ability of natural killer cells to lyse a sensitive target cell line. Lymphocyte proliferation can be stimulated with mitogens that bypass antigen specificity to activate cells or by stimulating the T cell receptor.

Pathways Between Stress and the Immune System

How could stress "get inside the body" to affect the immune response? First, sympathetic fibers descend from the brain into both primary (bone marrow and thymus) and secondary (spleen and lymph nodes) lymphoid tissues (Felten & Felten, 1994). These fibers can release a wide variety of substances that influence immune responses by binding to receptors on white blood

cells (Ader, Cohen, & Felten, 1995; Felten & Felten, 1994; Kemeny, Solomon, Morley, & Herbert, 1992; Rabin, 1999). Though all lymphocytes have adrenergic receptors, differential density and sensitivity of adrenergic receptors on lymphocytes may affect responsiveness to stress among cell subsets. For example, natural killer cells have both high-density and high-affinity β_2-adrenergic receptors, B cells have high density but lower affinity, and T cells have the lowest density (Anstead, Hunt, Carlson, & Burki, 1998; Landmann, 1992; Maisel, Fowler, Rearden, Motulsky, & Michel, 1989). Second, the hypothalamic–pituitary–adrenal axis, the sympathetic–adrenal–medullary axis, and the hypothalamic–pituitary–ovarian axis secrete the adrenal hormones epinephrine, norepinephrine, and cortisol; the pituitary hormones prolactin and growth hormone; and the brain peptides melatonin, β-endorphin, and enkephalin. These substances bind to specific receptors on white blood cells and have diverse regulatory effects on their distribution and function (Ader, Felten, & Cohen, 2001). Third, people's efforts to manage the demands of stressful experience sometimes lead them to engage in behaviors—such as alcohol use or changes in sleeping patterns—that also could modify immune system processes (Kiecolt-Glaser & Glaser, 1988). Thus, behavior represents a potentially important pathway linking stress with the immune system.

Maier and Watkins (1998) proposed an even closer relationship between stress and immune function: that the immunological changes associated with stress were adapted from the immunological changes in response to infection. Immunological activation in mammals results in a syndrome called *sickness behavior*, which consists of behavioral changes such as reduction in activity, social interaction, and sexual activity, as well as increased responsiveness to pain, anorexia, and depressed mood. This syndrome is probably adaptive in that it results in energy conservation at a time when such energy is best directed toward fighting infection. Maier and Watkins drew parallels between the behavioral, neuroendo-crine, and thermoregulatory responses to sickness and stress. The common thread between the two is the energy mobilization and redirection that is necessary to fight attackers both within and without.

Models of Stress, the Immune System, and Health

Conceptualizations of the nature of the relationship between stress and the immune system have changed over time. Selye's (1975) finding of thymic involution led to an initial model in which stress is broadly immunosuppressive. Early human studies supported this model, reporting that chronic forms of stress were accompanied by reduced natural killer cell cytotoxicity, suppressed lymphocyte proliferative responses, and blunted humoral responses to immunization (see S. Cohen, Miller, & Rabin, 2001; Herbert & Cohen, 1993;Kiecolt-Glaser, Glaser, Gravenstein, Malarkey, & Sheridan, 1996, for reviews). Diminished immune responses of this nature were assumed to be responsible for the heightened incidence of infectious and neoplastic diseases found among chronically stressed individuals (Andersen, Kiecolt-Glaser, & Glaser, 1994; S. Cohen & Williamson, 1991).

Although the global immunosuppression model enjoyed long popularity and continues to be influential, the broad decreases in immune function it predicts would not have been evolutionarily adaptive in life-threatening circumstances. Dhabhar and McEwen (1997, 2001) proposed that acute fight-or-flight stressors should instead cause redistribution of immune cells into the compartments in which they can act the most quickly and efficiently against invaders. In a series of experiments with mice, they found that during acute stress, T cells selectively redistributed into the skin, where they contributed to enhancement of the immune response. In contrast, during chronic stress, T cells were shunted away from the skin, and the immune response to skin test challenge was diminished (Dhabhar & McEwen, 1997). On the basis of these findings they proposed a biphasic model in which acute stress enhances, and chronic stress suppresses, the immune response.

A modification of this model posits that short-term changes in all components of the immune system (natural and specific) are unlikely to occur because they would expend too much energy to be adaptive in life-threatening circumstances. Instead, stress should shift the balance of the immune response toward activating natural processes and diminishing specific processes. The premise underlying this model is that natural immune responses are better suited to managing the potential complications of life-threatening situations than specific immune responses because they can unfold much more rapidly, are subject to fewer inhibitory constraints, and require less energy to be diverted from other bodily systems that support the fight-or-flight response (Dopp, Miller, Myers, & Fahey, 2000; Sapolsky, 1998).

Even with this modification of the biphasic model, neither it nor the global immunosuppression model sufficiently explains findings that link chronic stress with both disease outcomes associated with inadequate immunity (infectious and neoplastic disease) and disease outcomes associated with excessive immune activity (allergic and autoimmune disease). To resolve this paradox, some researchers have chosen to focus on how chronic stress might shift the balance of the immune response. The most well-known of these models hypothesizes that chronic stress elicits simultaneous enhancement and suppression of the immune response by altering patterns of cytokine secretion (Marshall et al., 1998). Th1 cytokines, which activate cellular immunity to provide defense against many kinds of infection and some kinds of neoplastic disease, are suppressed. This suppression has permissive effects on production of Th2 cytokines, which activate humoral immunity and exacerbate allergy and many kinds of autoimmune disease. This shift can occur via the effects of stress hormones such as cortisol (Chiappelli, Manfrini, Franceschi, Cossarizza, & Black, 1994). Th1-to-Th2 shift changes the balance of the immune response without necessarily changing the overall level of activation or function within the system. Because a diminished Th1-mediated cellular immune response could increase vulnerability to infectious and neoplastic disease, and an enhanced Th-2 mediated humoral immune response could increase vulnerability to autoimmune and allergic diseases, this cytokine shift model also is able to reconcile patterns of stress-related immune change with patterns of stress-related disease outcomes (Marshall et al., 1998).

Who Is Vulnerable to Stress-Induced Immune Changes?

If the stress response in the immune system evolved, a healthy organism should not be adversely affected by activation of this response because such an effect would likely have been selected against. Although there is direct evidence that stress-related immunosuppression can increase vulnerability to disease in animals (e.g., Ben Eliyahu, Shakhar, Page, Stefanski, & Shakhar, 2000; Quan et al., 2001; Shavit et al., 1985; Sheridan et al., 1998), there is little or no evidence linking stress-related immune change in healthy humans to disease vulnerability. Even large stress-induced immune changes can have small clinical consequences because of the redundancy of the immune system's components or because they do not persist for a sufficient duration to enhance disease susceptibility. In short, the immune system is remarkably flexible and capable of substantial change without compromising an otherwise healthy host.

However, the flexibility of the immune system can be compromised by age and disease. As humans age, the immune system becomes senescent (Boucher et al., 1998; Wikby, Johansson, Ferguson, & Olsson, 1994). As a consequence, older adults are less able to respond to vaccines and mount cellular immune responses, which in turn may contribute to early mortality (Ferguson, Wikby, Maxson, Olsson, & Johansson, 1995; Wayne, Rhyne, Garry, & Goodwin, 1990). The decreased ability of the immune system to respond to stimulation is one indicator of its loss of flexibility.

Loss of self-regulation is also characteristic of disease states. In autoimmune disease, for example, the immune system treats self-tissue as an invader, attacking it and causing pathology

such as multiple sclerosis, rheumatoid arthritis, Crohn's disease, and lupus. Immune reactions can also be exaggerated and pathological, as in asthma, and suggest loss of self-regulation. Finally, infection with HIV progressively incapacitates T-helper cells, leading to loss of the regulation usually provided by these cells. Although each of these diseases has distinct clinical consequences, the change in the immune system from flexible and balanced to inflexible and unbalanced suggests increased vulnerability to stress-related immune dysregulation; furthermore, dysregulation in the presence of disease may have clinical consequences (e.g., Bower, Kemeny, Taylor, & Fahey, 1998).

The Present Analysis

We performed a meta-analysis of published results linking stress and the immune system. We feel that this area is in particular need of a quantitative review because of the methodological nature of most studies in this area. For practical and economic reasons, many psychoneuroimmunology studies have a relatively small sample size, creating the possibility of Type II error. Furthermore, many studies examine a broad range of immunological parameters, creating the possibility of Type I error. A quantitative review, of which meta-analysis is the best example, can better distinguish reliable effects from those arising from both Type I and Type II error than can a qualitative review.

We combined studies in such a way as to test the models of stress and immune change reviewed above. First, we examined each stressor type separately, yielding separate effects for stressors of different duration and trajectory. Second, we examined both healthy and medical populations, allowing comparison of the effects of stress on resilient and vulnerable populations; along the same lines, we also examined the effects of age. Finally, we examined all immune parameters separately so that patterns of response (e.g., global immunosuppression vs. cytokine shift) would be clearer.

Method

Article Identification

Articles for the meta-analysis were identified through computerized literature searches and searches of reference lists. MEDLINE and PsycINFO were searched for the years 1960 –2001. Following the example of Herbert and Cohen (1993), we used the terms *stress*, *hassles*, and *life events* in combination with the term *immune* to search both databases. The reference lists of 11 review articles on stress and the immune system (Benschop, Geenen, et al., 1998; Biondi, 2001; Cacioppo, 1994; S. Cohen & Herbert, 1996; S. Cohen et al., 2001; Herbert & Cohen, 1993; Kiecolt-Glaser, Cacioppo, Malarkey, & Glaser, 1992; Kiecolt-Glaser, McGuire, Robles, & Glaser, 2002; Maier, Watkins, & Fleshner, 1994; O'Leary, 1990; Zorrilla et al., 2001) were then searched to identify additional articles.

We selected only articles that met a number of inclusion criteria. The first criterion was that the work had to include a measure of stress. This criterion could be met if a sample experiencing a stressor was compared with an unstressed control group, if a sample experiencing a stressor was compared with itself at a baseline that could reasonably be considered low stress, or if differing degrees of stress in a sample were assessed with an explicit measure of stress. This criterion was not met if, for example, anxiety—an affective state—was used as a proxy for stress, or it seemed likely that a "baseline" assessment occurred during periods of significant stress. The second criterion was that the stressor had to be psychosocial. Stressors that included a significant physical element such as pain, cold, or physical exhaustion were eliminated (e.g., Antarctic isolation, space flight, military training). The third criterion was that the work had to include a measure of the immune system. This criterion was met by any enumerative or functional in vitro or in vivo immune assay. However, clinical disease outcomes such as HIV

progression or rhinovirus infection did not meet this criterion. Finally, we eliminated articles from which a meaningful effect size could not be abstracted. For example, when between- and within-subjects observations were treated as independent, the reported effect was likely to be inflated. In a few cases, effects of stress and clinical status were confounded—that is, a stressed clinical group was compared with an unstressed healthy group—and hence these studies were excluded from the meta-analysis.

Stressor Classification

We coded stressors in the articles into five classes: acute time-limited, brief naturalistic, event sequence, chronic, and distant. The most difficult distinctions among event sequence, chronic, and distant stressors were based on temporal and qualitative considerations. Event sequences included discrete stressors occurring 1 year or less before immune assessment and could be of any severity. These were most often normative stressors such as bereavement. Chronic stressors were ongoing stressors such as caregiving and disability. Distant stressors were severe, traumatic events that could meet the stressor criterion for posttraumatic stress disorder (American Psychiatric Association, 1994), such as combat exposure or abuse, and had happened more than 1 year before immune assessment. Most stressors in this category occurred 5 to 10 years before immune assessment. Disagreements in stressor classification were resolved by consensus. Subgroups for moderator analyses were similarly decided.

The Meta-Analysis

Overview of procedures—Meta-analysis is a tool for synthesizing research findings. It proceeds in two phases. In the first, effect sizes are computed for each study. An effect size represents the magnitude of the relationship between two variables, independent of sample size. In this context it can be viewed as a measure of how much two groups, one experiencing a stressor and the other not, differ on a specific immune outcome. In the second phase, effect sizes from individual studies are combined to arrive at an aggregate effect size for each immune outcome of interest.

We used Pearson's *r* as the effect size metric in this meta-analysis. Effect sizes for individual studies were computed using descriptive statistics presented in the original published reports. When these statistics were not available, we requested them from authors. This strategy was successful in most circumstances. To compute Pearson's *r* from descriptive statistics in between-subjects designs, we subtracted the control group mean from the stressed group mean and divided this value by the pooled sample standard deviation. The value that emerged from this computation, known as Cohen's *d*, was then converted into a Pearson's *r* by taking the square root of the quantity $d^2/(d^2 + 4)$. (See Rosenthal, 1994.) To compute Pearson's *r* from descriptive statistics in within-subjects designs, we subtracted the group mean at baseline from the group mean during stress and divided this quantity by the sample standard deviation at baseline. This *d* value was converted into a Pearson's *r* by taking the square root of the quantity $d^2/(d^2 + 4)$. In cases in which descriptive statistics were not available, Pearson's *r* was computed from inferential statistics using standard formulae (Rosenthal, 1994). These formulae had to be modified slightly for studies that used within-subjects designs because effect sizes are systematically overestimated when they are calculated from repeated measures test statistics (Dunlap, Cortina, Vaslow, & Burke, 1996). In these situations we derived effect size estimates using the formula $d = t_c[2\ (1 - r)]^{1/2}$, where t_c corresponds to the value of the *t* statistic for correlated measures, and *r* corresponds to the value of the correlation between outcome measures at pretest and posttest (Dunlap et al., 1996). Because very few studies reported the value of *r*, we used a value of .60 to compute effect sizes in this meta-analysis. This represents the average correlation between pre-stress and poststress measures of immune function in a series of studies performed in our laboratories. To ensure that the meta-analytic findings were robust to variations in *r*, we conducted follow-up analyses using *r* values ranging from .45 to .

75. Very similar findings emerged from these analyses, suggesting that the values we present below are reliable estimates of effect size. If anything, they are probably conservative estimates, because the pre–post correlation between immune measures often is substantially lower than .60.

The effect size estimates from individual studies were subsequently aggregated using random-effects models with the software program Comprehensive Meta-Analysis (Borenstein & Rothstein, 1999). The random-effects model views each study in a meta-analysis as a random observation drawn from a universe of potential investigations. As such, it assumes that the magnitude of the relationship between stress and the immune system differs across studies as a result of random variance associated with sampling error and differences across individuals in the processes of interest. Because of these assumptions, random-effects models not only permit one to draw inferences about studies that have been done but also to generalize to studies that might be done in the future (Raudenbush, 1994; Shadish & Haddock, 1994). It also bears noting that in the population of studies on stress and immunity there is likely to be a fair amount of nonrandom variance, as researchers who examine ostensibly similar phenomena may still differ in terms of the samples they recruit, the operational definition of stress they use, and the laboratory methods they utilize to assess a specific immune process.

Separate random-effects models were computed for each immune outcome included in the meta-analysis. Prior to computing the random-effects model, *r* values derived from each study were *z*-transformed by the software program, as recommended by Shadish and Haddock (1994), to stabilize variance. The *z* values were later back-transformed into *r* values to facilitate interpretation of the meta-analytic findings. In the end, each random-effects model yielded an aggregate weighted effect size *r*, which can be interpreted the same way as a correlation coefficient, ranging in value from −1.00 to 1.00. Each *r* statistic was weighted before aggregation by multiplying its value by the inverse of its variance; this procedure enabled larger studies to contribute to effect size estimates to a greater extent than smaller ones. Weighting effect sizes is important because larger studies provide more accurate estimates of true population parameters (Shadish & Haddock, 1994). After each aggregate effect size had been derived, we computed 95% confidence intervals around it, assessed whether it was statistically significant, and computed a heterogeneity coefficient to determine whether the studies contributing to it had yielded consistent findings. Following convention, aggregate effect sizes were considered statistically different from zero when (a) their corresponding *z* value was greater than 1.96 and (b) the 95% confidence intervals around them did not include the value zero (Rosenthal, 1991; Shadish & Haddock, 1994).

To determine whether the studies contributing to each aggregate effect size shared a common population value, we computed the heterogeneity statistic *Q* (Shadish & Haddock, 1994). This statistic is chi-square distributed with $k - 1$ degrees of freedom, where *k* represents the number of independent effect sizes included. When a statistically significant heterogeneity test emerged, we searched for moderators (characteristics of the participants, stressful experience, or measurement strategy) that could explain the variability across studies. The first step in this process involved estimating correlations between participant characteristics (e.g., mean age, percentage female) and immune effects to examine whether the strength of effects varied according to demographics. When it was possible to do so, we then stratified the studies according to characteristics of the stressful experience (e.g., duration, quality) or the measurement strategy (e.g., interview, checklist), and computed separate random-effects analyses for each subgroup.

Handling missing data—Occasionally authors of studies failed to report the descriptive or inferential statistics needed to compute an effect size. In some of these cases, the authors noted that there was a significant difference between a stressed and control group. When this

occurred, we computed effect sizes assuming that *p* values were equivalent to .05. This represents a conservative approach because the actual *p* values were probably smaller. In other cases, the authors noted that a stressed and control group did not differ with respect to an immune outcome, but failed to provide any further statistical information. When this occurred, we computed effect sizes assuming that there was no difference at all between the groups (*r* = .00). Because there is seldom no difference at all between two groups, this also represents a conservative strategy. Imputation was used in less than 7% of cases.

Handling dependent data—The validity of a meta-analysis rests on the assumption that each value contributing an aggregate effect size is statistically independent of the others (Rosenthal, 1991). We devised a number of strategies to avoid violating this independence assumption. First, in studies that assessed stimulated-lymphocyte proliferation at multiple mitogen dosages, we computed the average effect size across mitogen dosages, and we used this value to derive aggregate indices. We used an analogous strategy for studies that assessed natural killer cell cytotoxicity at multiple effector:target cell ratios. Second, in studies that utilized designs in which multiple laboratory stressors were compared with a control condition, the average effect size across stressor conditions was computed and later used to derive aggregate indices. Because this averaging procedure in most cases yielded an effect size that was smaller than that of the most potent stressor, we also computed meta-analyses using the larger of the effect sizes from each study rather than the average. Doing so did not alter any of the substantive findings we report. Third, in studies in which immune outcomes were assessed on multiple occasions during a stressful experience, the average effect size across occasions was used to derive aggregate indices. Note that we did not conduct meta-analyses of recovery effects, that is, immune values after a stressor had ended. Although such an analysis would answer interesting questions about the stress-recovery process, there were not enough studies that included similar immune outcomes assessed at similar time points after stress to permit a complete analysis. Fourth, because some data were published in more than one outlet, we contacted authors of multiple publications to determine sample independence or dependence.

Results

Preliminary Findings

The meta-analysis is based on effect sizes derived from 293 independent studies. These studies were reported in 319 separate articles in peer-reviewed scientific journals (see Table 2). A total of 18,941 individuals participated in these studies. Their mean age was 34.8 years (*SD* = 15.9). Although the studies collectively included a broad range of age groups (range = 5–78 years), most focused heavily on younger adults. More than half of the studies (51.3%) had a mean age under 30.0 years, and more than four fifths (84.8%) had a mean age under 55.0 years. Slightly more than two thirds of the studies (68.5%) included women; in the average study almost half (42.8%) of the participants were female. The vast majority of studies (84.8%) focused on medically healthy adults.[2] Of those that included medical populations, most focused on HIV/AIDS (*k* = 18; 38.3%), arthritis (*k* = 6; 12.8%), cancer (*k* = 5; 10.6%), or asthma (*k* = 4; 8.5%).

With respect to the kinds of stressors examined by studies in the meta-analysis, the most commonly utilized models were acute laboratory challenges (*k* = 85; 29.0%) and brief naturalistic stressors (*k* = 63; 21.5%). Stressful event sequences (*k* = 30; 10.2%), chronic

[2]The proportion of student samples varied across stressor categories. Nearly all of the studies of brief naturalistic stressors used student samples (*k* = 60; 95.2%) because these stressors were predominantly examinations. Student samples were also used in a large minority of acute time-limited stressor studies (*k* = 31; 40.5%) but constituted a small minority of samples used in studies of life-event checklists (*k* = 8; 14.0%) and studies of event sequences (*k* = 2; 6.6%), and student samples were not used in studies of chronic stressors or stress appraisals and intrusions. These are rough estimates, as some studies did not specify whether young adult samples were drawn from a student population.

stressors (k = 23; 7.8%), and distant traumatic experiences (k = 9; 3.1%) were explored less frequently. More than a quarter of the studies in the meta-analysis modeled the stress process by administering nonspecific life-event checklists (k = 53; 18.1%) and/or global perceived stress measures (k = 21; 7.1%) to participants. A small minority of studies examined whether reports of perceived stress or intrusive memories were associated with the extent of immune dysregulation within populations who had suffered a specific traumatic experience (k = 9; 3.1%).

The studies in the meta-analysis examined 292 distinct immune system outcomes. A minority of these outcomes were assessed in three or more studies (k = 87; 30.0%), and as such, they are the focus of the meta-analyses we present in the rest of this article (see Table 1). The most commonly assessed enumerative outcomes were counts of T-helper lymphocytes (k = 90; 30.7%), T-cytotoxic lymphocytes (k = 81; 27.6%), natural killer cells (k = 67; 22.9%), and total lymphocytes (k = 52; 17.7%). The most commonly assessed functional outcomes were natural killer cell cytotoxicity (k = 94; 32.1%) and lymphocyte proliferation stimulated by the mitogens phytohemagglutinin (PHA; k = 65; 22.2%), concanavalin A (ConA; k = 39; 13.3%), and pokeweed mitogen (PWM; k = 26; 8.9%).

Interpreting the Meta-Analytic Findings

Table 1 lists the immune parameters analyzed with the arm of the immune system to which they belong (natural or specific) and, briefly, their function. Where relevant, cell surface markers used to identify classes of immunocytes in flow cytometry are given. For example, the cell surface marker CD19 is used to identify B lymphocytes. Recall that different models of stress and the immune system posit differential effects of stress on subsets of the immune system—for example, natural versus specific immunity or cellular (Th1) versus humoral (Th2) immunity. Table 1 acts as a guide for interpreting the pattern of results in light of these models.

In the following sections we describe the meta-analytic results for each stressor category. A useful rule of thumb for judging effect sizes is to consider values of .10, .30, and .50 as corresponding to small, medium, and large effects, respectively (J. Cohen & Cohen, 1983); more generally, the aggregate effect size r can be interpreted in the same fashion as a correlation, with values ranging from −1.00 to 1.00. Positive values indicate that the presence of a stressor increases a particular immune parameter relative to some baseline (or control) condition. We should caution the reader that in some analyses, our statistics are derived from as few as three independent studies. Although meta-analyses of small numbers of studies do not pose any major statistical problems, it is important to remember that they have limited power to detect statistically significant effect sizes. What a meta-analysis can accurately provide in these instances, however, is an estimate of how much and what direction a given stressor's presence influences a specific immune outcome (i.e., an effect size estimate).

Meta-Analytic Results for the Effects of Stressors

Acute time-limited stressors—Acute time-limited stressors included primarily experimental manipulations of stressful experiences, such as public speaking and mental arithmetic, that lasted between 5 and 100 min. Reliable effects on the immune system included increases in immune parameters, especially natural immunity. The most robust effect of this kind of experience was a marked increase in the number of natural killer cells (r =.43) and large granular lymphocytes (r =.53) in peripheral blood (see Table 3). This effect is consistent with the view that acute stressors cause immune cells to redistribute into the compartments in which they will be most effective (Dhabhar & McEwen, 1997). However, other types of lymphocytes did not show robust redistribution effects: B cells and T-helper cells showed very little change (rs = −.07 and .01, respectively), and this change was not statistically significant across studies. T-cytotoxic lymphocytes did tend to increase reliably in peripheral blood,

though to a lesser degree than their natural immunity counterparts (r =.20); this increase drove a reliable decline in the T-helper:T-cytotoxic ratio (r = −.23). However, natural killer cells as well as T-cytotoxic cells can express CD8, the marker most often used to define the latter population. Because some studies did not use the T cell receptor (CD3) to differentiate between CD3–CD8+ natural killer cells and CD3+CD8+ T-cytotoxic cells, it is possible that the effect for "T-cytotoxic cells" is actually being driven by natural killer cells (Benschop, Rodriguez-Feuerhahn, & Schedlowski, 1996).

The results for cell percentages roughly parallel those for number. However, the percentage data are harder to interpret because any given parameter is linearly dependent on the other parameters: For example, the enumerative data suggest that the decrease in percentage T-helper cells (r = −.24) is probably an artifact of the increases in percentage natural killer cells (r = .24) and percentage T-cytotoxic cells (r = .09).

Another effect that may be considered a redistribution effect is the significant increase in secretory IgA in saliva (r = .22). The time frame of these acute stressors is too short for the synthesis of a significant amount of new antibody; therefore, this increase is probably due to release of already-synthesized antibody from plasma cells and increased translocation of antibody across the epithelium and into saliva (Bosch, Ring, de Geus, Veerman, & Amerongen, 2002). This effect therefore represents relocation, albeit of an immune protein rather than an immune cell.

There were also a number of functional effects. First, natural killer cell cytotoxicity significantly increased with acute stressors (r = .30), but only when the concomitant increase in proportion of natural killer cells in the effector mix was not removed statistically. When examined on a per-cell basis, cytotoxicity did not significantly increase (r = .12). One could, therefore, consider the increase in cytotoxicity a methodological artifact of the definition of *effector* in effector:target ratios. However, to the degree that one is interested in the general cytotoxic potential of the contents of peripheral blood rather than that of a specific natural killer cell, the uncorrected value is more illustrative. Second, mitogen-stimulated proliferative responses decreased significantly. Again, this could be a methodological artifact of the mix of cells in the assay. However, the proportion of total T and B cells, which are responsible for the proliferative response to PWM and ConA, did not decrease as reliably or as much as did the proliferative response (rs = −.05 to −.11 vs. −.10 to −.17), suggesting that acute stressors do decrease this function of specific immunity. Finally, the production of two cytokines, IL-6 and IFNγ, was increased significantly following acute stress (rs = .28 and .21, respectively).

The data for acute stressors, therefore, support an upregulation of natural immunity, as reflected by increased number of natural killer cells in peripheral blood, and potential downregulation of specific immunity, as reflected by decreased proliferative responses. Other indicators of upregulated natural immunity include increased neutrophil numbers in peripheral blood (r = .30), increased production of a proinflammatory cytokine (IL-6), and increased production of a cytokine that potently stimulates macrophages and natural killer cells as well as T cells (IFNγ). The only exception to this pattern was the increased secretion of IgA antibody, which is a product of the specific immune response. An interesting question for future research is whether this effect is part of a larger nonspecific protein release in the oral cavity in response to acute stress (cf. Bosch et al., 2002).

It bears noting that a number of the findings presented in Table 3 are accompanied by significant heterogeneity statistics. To identify moderating variables that might explain some of this heterogeneity, we examined whether effect sizes varied according to demographic characteristics of the sample (mean age and percentage female) or features of the acute challenge (its duration and nature). Neither of the demographic characteristics showed a

consistent relationship with immune outcomes. Although these findings suggest that acute time-limited stressors elicit a similar pattern of immune response for men and women across the life span, this conclusion needs to be viewed somewhat cautiously given the narrow range of ages found in these studies. We also did not find a consistent pattern of relationships between features of the acute challenge and immune outcomes. Acute stressors elicited similar patterns of immune change across a wide spectrum of durations ranging from 5 though 100 min and irrespective of whether they involved social (e.g., public speaking), cognitive (e.g., mental arithmetic), or experiential (e.g., parachute jumping) forms of stressful experience.

Brief naturalistic stressors—Table 4 presents the meta-analysis of brief naturalistic stressors for medically healthy adults. The vast majority of these stressors ($k = 60$; 95.2%) involved students facing academic examinations. In contrast to the acute time-limited stressors, examination stress did not markedly affect the number or percentage of cells in peripheral blood. Instead, the largest effects were on functional parameters, particularly changes in cytokine production that indicate a shift away from cellular immunity (Th1) and toward humoral immunity (Th2). Brief stressors reliably changed the profile of cytokine production via a decrease in a Th1-type cytokine, IFNγ ($r = -.30$), which stimulates natural and cellular immune functions, and increases in the Th2-type cytokines IL-6 ($r = .26$), which stimulates natural and humoral immune functions, and IL-10 ($r = .41$), which inhibits Th1 cytokine production. Note that IFNγ and IL-6 share the property of stimulating natural immunity but differentially stimulate cytotoxic versus inflammatory effector mechanisms. Their dissociation after brief naturalistic stress indicates differential effects between Th1 and Th2 responses rather than natural and specific responses.

The functional assay data are consistent with this suggestion of suppression of cellular immunity via decreased Th1 cytokine production: The T cell proliferative response significantly decreased with brief stressors ($r = -.19$ to $-.32$), as did natural killer cell cytotoxicity ($r = -.11$). Increased antibody production to latent virus, particularly Epstein-Barr virus ($r = .20$), is also consistent with suppression of cellular immunity, enhancement of humoral immunity, or both.

There was also evidence that age contributed to vulnerability to stress-related immune change during brief naturalistic stressors, even within a limited range of relatively young ages. When we examined whether effect sizes varied according to demographic characteristics of the sample, sex ratio did not show a consistent pattern of relations with immune processes. However, the mean age of the sample was strongly related to study effect size. To the extent that a study enrolled participants of older ages, it was likely to observe more pronounced decreases in natural killer cell cytotoxicity ($r = -.58$, $p = .04$; $k = 14$), T lymphocyte proliferation to the mitogens PHA ($r = -.58$, $p = .04$; $k = 13$) and ConA ($r = -.31$, $p = .38$; $k = 9$), and production of the cytokine IFNγ ($r = -.63$, $p = .09$; $k = 8$) in response to brief naturalistic stress. The strength of these findings is particularly surprising given the narrow range of ages found in studies of brief natural stress; the mean participant age in this literature ranged from 15.7 to 35.0 years.

We also calculated effect sizes for three studies examining the effects of examination stress on individuals with asthma (see Table 5). These three studies, all emanating from a team of investigators at the University of Wisconsin—Madison, found that stress reliably increased superoxide release ($r = .20$ to $.37$) and decreased natural killer cell cytotoxicity ($r = -.33$). Because natural killer cells are stimulated by Th1 cytokines, this change is consistent with a Th1-to-Th2 shift. However, stress also reliably increased T cell proliferation to PHA ($r = .32$), which is not consistent with such a shift. The generally larger effect sizes are consistent with the idea that individuals with immunologically mediated disease are more susceptible to stress-related immune dysregulation, but the reversed sign for T cell proliferation also indicates that

that pattern of dysregulation may also be more disorganized. That is, the organized pattern of suppression of Th1 but not Th2 immune responses in healthy individuals undergoing brief stressors may reflect regulation in the healthy immune system. In contrast, the lack of regulation in a diseased immune system may lead to more chaotic changes during stressors.

Stressful event sequences—The meta-analysis of stressful event sequences is presented in Table 6. With the exception of significant increases in the number of circulating natural killer cells and the number of antibodies to the latent Epstein-Barr virus, the findings indicate that stressful event sequences are not associated with reliable immune changes. For many immune outcomes, however, significant heterogeneity statistics are evident. Studies of healthy adults generally fell into two categories that yielded disparate patterns of immune findings. The largest group of studies focused on the death of a spouse as a stressor and, as such, used samples consisting primarily of older women. Collectively, these studies found that losing a spouse was associated with a reliable decline in natural killer cell cytotoxicity ($r = -.23$, $p = .01$; $k = 6$) but not with alterations in stimulated-lymphocyte proliferation by the mitogens ConA ($r = -.04$, $p = .45$; $k = 4$), PHA ($r = -.01$, $p = .93$; $k = 7$), or PWM ($r = -.08$, $p = .76$; $k = 3$) or with changes in the number of T-helper lymphocytes ($r = .07$, $p = .52$; $k = 6$) or T-cytotoxic lymphocytes ($r = -.13$, $p = .45$; $k = 5$) in peripheral blood. The next largest group of studies in this area examined immune responses to disasters, which may have different neuroendocrine consequences than loss; whereas loss is generally associated with increases in cortisol, trauma may be associated with decreases in cortisol (Yehuda, 2001;Yehuda, McFarlane, & Shalev, 1998). Natural disaster samples tended to focus on middle-aged adults of both sexes who were direct victims of the disaster, rescue workers at the scene, or personnel at nearby medical centers. There were medium-size effects suggesting increases in natural killer cell cytotoxicity ($r = .25$, $p = .53$; $k = 4$) and stimulated-lymphocyte proliferation by the mitogen PHA ($r = .26$, $p = .33$; $k = 2$), as well as decreases in the number of T-helper lymphocytes ($r = -.20$, $p = .43$; $k = 2$) and T-cytotoxic lymphocytes ($r = -.23$, $p = .55$; $k = 2$) in the circulation. However, none of them was statistically significant because of the small number of studies involved, and therefore these effects should be considered suggestive but not reliable.

An additional group of studies in this area examined immune responses to a positive initial biopsy for breast cancer in primarily middle-aged female participants before and after the procedure. The three studies of this nature did not yield a consistent pattern of relations with any of the immune outcomes.

In summary, stressful event sequences did not elicit a robust pattern of immune changes when considered as a whole. When these sequences are broken down into categories reflecting the stressor's nature, the meta-analysis yields evidence of declines in natural immune response following the loss of a spouse, nonsignificant increases in natural and specific immune responses following exposure to natural disaster, and no immune alterations with breast biopsy. Unfortunately, we cannot determine whether these disparate patterns of immune response are attributable to features of the stressors, demographic or medical characteristics of the participants, or some interaction between these factors.

Chronic stressors—Chronic stressors included dementia caregiving, living with a handicap, and unemployment. Like other nonacute stressors, they did not have any systematic relationship with enumerative measures of the immune system. They did, however, have negative effects on almost all functional measures of the immune system (see Table 7). Both natural and specific immunity were negatively affected, as were Th1 (e.g., T cell proliferative responses) and Th2 (e.g., antibody to influenza vaccine) parameters. The only nonsignificant change was for antibody to latent virus; this effect size was substantial ($r = .44$), but there was also substantial heterogeneity. Further analyses showed that demographics did not moderate

this effect: Immune responses to chronic stressors were equally strong across the age spectrum as well as across sex.

Distant stressors—Distant stressors were traumatic events such as combat exposure or abuse occurring years prior to immune assessment. The meta-analytic results for distant stressors appear in Table 8. The only immune outcome that has been examined regularly in this literature is natural killer cell cytotoxicity, and it is not reliably altered in persons who report a distant traumatic experience.

Meta-Analytic Results for the Effects of Checklists and Ratings

Nonspecific life events—Most of the studies in this area examined whether immune responses varied as a function of the number of life events a person endorsed on a standard checklist, a person's rating of the impact of those events, or both. As Table 9 illustrates, this methodology yielded little in the way of significant outcomes in healthy participants. To determine whether vulnerability to life events might vary across the life span, we divided studies into two categories on the basis of a natural break in the age distribution. These analyses provided evidence that older adults are especially vulnerable to life-event–induced immune change. In studies that used samples of adults who had a mean age above 55, life events were associated with reliable declines in lymphocyte-proliferative responses to PHA ($r = -.40$, $p = .05$; $k = 2$) and natural killer cell cytotoxicity ($r = -.59$, $p = .001$; $k = 2$). These effects were much weaker in studies with a mean age below 55: Life events were not associated with proliferative responses to PHA ($r = -.22$, $p = .24$; $k = 2$), and showed a reliable but modest relationship with natural killer cell cytotoxicity ($r = -.10$, $p = .03$; $k = 8$). The differences in effect size between older and younger adults were statistically significant for natural killer cell cytotoxicity ($p < .001$) but not PHA-induced proliferation ($p < .15$). None of the other moderators we examined—sex ratio, kind of life event assessed (daily hassle vs. major event), or the method used to do so (checklist vs. interview)—was related to immune outcomes.

Table 10 presents the relationship between life events and immune parameters in participants with HIV/AIDS. The presence of life events was associated with a significant reduction in the number of natural killer cells and a marginal reduction in the number of T-cytotoxic lymphocytes. It is unrelated to the number of T-helper lymphocytes, the percentage of T-cytotoxic lymphocytes, and the T-helper:T-cytotoxic ratio, all of which are recognized indicators of disease progression for patients with HIV/AIDS.

We have already proposed that immunological disease diminishes the resilience and self-regulation of the immune system, making it more vulnerable to stress-related disruption, and this may be the case in HIV-infected versus healthy populations. However, studies of HIV-infected populations also utilized more refined measures of life events (interviews that factor in biographical context) than did studies of healthy populations (typically, checklist measures). Unfortunately, we cannot differentiate between these explanations on the basis of the available data.

Global stress appraisals and intrusive thoughts—The meta-analysis of stress appraisals and intrusive thoughts is displayed in Table 11. These studies generally enrolled large populations of adults who were not experiencing any specific form of stress and examined whether their immune responses varied according to stress appraisals and/or intrusive thoughts. This methodology was unsuccessful at documenting immune changes related to stress. Because of the small number of studies in this category, moderator analyses could not be performed.

The meta-analysis results shown in Table 12 address a similar question with regard to persons who are in the midst of a specific event sequence or a chronic stressor. To the extent that they appraise their lives as stressful or report the occurrence of intrusive thoughts, these individuals

exhibit a significant reduction in natural killer cell cytotoxicity. Although this effect does not extend to the number of T-helper and T-cytotoxic lymphocytes in the circulation, it suggests that a person's subjective representation of a stressor may be a determinant of its impact on the immune response.

Evidence Regarding Type I Error and Publication Bias

The large number of effect sizes generated by the meta-analysis raises the possibility of Type I error. One strategy for evaluating this concern involves dividing the number of significant findings in a meta-analysis by the total number of analyses conducted. When we performed this calculation, a value of 25.6% emerged, suggesting that more than one fourth of the analyses yielded reliable findings. This exceeds the 5% value at which investigators typically become concerned about Type I error rates and gives us confidence that the meta-analytic findings presented here are robust.

A second concern arises from the publication bias toward positive findings, which could skew meta-analytic results toward larger effect sizes. Fortunately, recent advances in meta-analysis enable one to evaluate the extent of this publication bias by using graphical techniques. A funnel plot can be drawn in which effect sizes are plotted against sample sizes for any group of studies. Because most studies in any given area have small sample sizes and therefore tend to yield more variable findings, the plot should end up looking like a funnel, with a narrow top and a wide bottom. If there is a bias against negative findings in an area, the plot is shifted toward positive values or a chunk of it will be missing entirely.

We drew funnel plots for all of the immune outcomes in the meta-analysis for which there were a sufficient number of observations. Although not all of them yielded perfect funnels, there was no systematic evidence of publication bias. Space limitations prevent us from including all plots; however, Figure 1 displays three plots that are prototypical of those we drew. As is evident from the data in the figure, psychoneuroimmunology researchers seem to be reporting positive and negative findings—and not hiding unfavorable outcomes when they do emerge. Thus, we do not have any major concerns about publication bias leading this meta-analysis to dramatically overestimate effect sizes.

Discussion

The immune system, once thought to be autonomous, is now known to respond to signals from many other systems in the body, particularly the nervous system and the endocrine system. As a consequence, environmental events to which the nervous system and endocrine system respond can also elicit responses from the immune system. The results of meta-analysis of the hundreds of research reports generated by this hypothesis indicate that stressful events reliably associate with changes in the immune system and that characteristics of those events are important in determining the kind of change that occurs.

Models of Stress and the Immune System

Selye's (1975) seminal findings suggested that stress globally suppressed the immune system and provided the first model for how stress and immunity are related. This model has recently been challenged by views that relations between stress and the immune system should be adaptive, at least within the context of fight-or-flight stressors, and an even newer focus on the balance between cellular and humoral immunity. The present meta-analytic results support three of these models. Depending on the time frame, stressors triggered adaptive upregulation of natural immunity and suppression of specific immunity (acute time-limited), cytokine shift (brief naturalistic), or global immunosuppression (chronic).

When stressors were acute and time-limited—that is, they generally followed the temporal parameters of fight-or-flight stressors—there was evidence for adaptive redistribution of cells and preparation of the natural immune system for possible infection, injury, or both. In evolution, stressor-related changes in the immune system that prepared the organisms for infections resulting from bites, puncture wounds, scrapes, or other challenges to the integrity of the skin and blood could be selected for. This process would be most adaptive when it was also efficient and did not divert excess energy from fight-or-flight behavior. Indeed, changes in the immune system following acute stress conformed to this pattern of efficiency and energy conservation. Acute stress upregu-lated parameters of natural immunity, the branch of the immune system in which most changes occurred, which requires only minimal time and energy investment to act against invaders and is also subject to the fewest inhibitory constraints on acting quickly (Dopp et al., 2000; Sapolsky, 1998). In contrast, energy may actually be directed away from the specific immune response, as indexed by the decrease in the proliferative response. The specific immune response in general and proliferation in particular demand time and energy; therefore, this decrease might indicate a redirection away from this function. Similar redirection occurs during fight-or-flight stressors with regard to other nonessential, future-oriented processes such as digestion and reproduction. As stressors became more chronic, the potential adaptiveness of the immune changes decreased. The effect of brief stressors such as examinations was to change the potency of different arms of specific immunity —specifically, to switch away from cellular (Th1) immunity and toward humoral (Th2) immunity.

The stressful event sequences tended to fall into two substantive groups: bereavement and trauma. Bereavement was associated with decreased natural killer cell cytotoxicity. Trauma was associated with nonsignificantly increased cytotoxicity and increased proliferation but decreased numbers of T cells in peripheral blood. The different results for loss and trauma mirror neuroendocrine effects of these two types of adverse events. Loss—maternal separation in nonhuman animals and bereavement in humans—is commonly associated with increased cortisol production (Irwin, Daniels, Risch, Bloom, & Weiner, 1988; Laudenslager, 1988; McCleery, Bhagwagar, Smith, Goodwin, & Cowen, 2000). In contrast, trauma and posttraumatic stress disorder are commonly associated with decreased cortisol production (see Yehuda, 2001; Yehuda et al., 1998, for reviews). To the degree that cortisol suppresses immune function such as natural killer cell cytotoxicity, these results have the potential to explain the different effects of loss and trauma event sequences.

The most chronic stressors were associated with the most global immunosuppression, as they were associated with reliable decreases in almost all functional immune measures examined. Increasing stressor duration, therefore, resulted in a shift from potentially adaptive changes to potentially detrimental changes, initially in cellular immunity and then in immune function more broadly. It is important to recognize that although the effects of chronic stressors may be due to their duration, the most chronic stressors were associated with changes in identity or social roles (e.g., acquiring the role of caregiver or refugee or losing the role of employee). These chronic stressors may also be more persistent, that is, constantly rather than intermittently present. Finally, chronic stressors may be less controllable and afford less hope for control in the future. These qualities could contribute to the severity of the stressor in terms of both its psychological and physiological impact.

Increasing stressor chronicity also impacted the type of parameter in which changes were seen. Compared with the natural immune system, the specific immune system is time and energy intensive and as such is expected to be invoked only when circumstances (either a stressor or an infection; cf. Maier & Watkins, 1998) persist for a longer period of time. Affected immune domains—natural versus specific—were consistent with the duration of the stressors—acute versus chronic. Furthermore, changing immune responses via redistribution of cells can happen

much faster than changes via the function of cells. The time frames of the stressor and the immune domain were also consistent; acute stress affected primarily enumerative measures, whereas stressors of longer duration affected primarily functional measures.

The results of these analyses suggest that the dichotomization of the immune system into natural and specific categories and, within specific immunity, into cellular and humoral measures, is a useful starting point with regard to understanding the effects of stressors. Categorizing an immune response is a difficult process, as each immune response is highly redundant and includes natural, specific, cellular, and humoral immune responses acting together. Given this redundancy, the differential results within these theoretical divisions were remarkably, albeit not totally, consistent. As further immunological research defines these divisions more subtly, the results with regard to stressors may become even clearer. However, the present results suggest that the categories used here are meaningful.

The results of this meta-analysis reflect the theoretical and empirical progress of this literature over the past 4 decades. Increased differentiation in the quality of stressors and the immunological parameters investigated have allowed complex models to be tested. In contrast, previous meta-analyses were bound by a small number of more homogenous studies. Herbert and Cohen (1993) reported on 36 studies published between 1977 and 1991, finding broadly immunosuppressive effects of stress. Zorrilla et al. (2001) reported on 82 studies published between 1980 and 1996, finding potentially adaptive effects of acute stressors in addition to evidence for immunosuppression with longer stressors. It is important to note that meta-analytic findings are bound by the models tested in the literature. As more complex models are tested, more complex relationships emerge in meta-analysis. We next consider some such areas of complexity that should be considered in future psychoneuroimmunology research.

Individual Differences and Immune Change Under Stress

The meta-analytic results indicate that organismic variables such as age and disease status moderate vulnerability to stress-related decreases in functional immune measures. Both aging and HIV are associated with immune senescence and loss of responsiveness (Effros et al., 1994; Effros & Pawelec, 1997), and both are also associated with disruption of neuroendocrine inputs to the immune system (Kumar et al., 2002; Madden, Thyagarajan, & Felten, 1998). The loss of self-regulation in disease and aging likely makes affected people more susceptible to negative immunological effects of stress. Finally, the meta-analysis did not reveal effects of sex on immune responses to stressors. However, these comparisons simply correlated the sex ratio of the studies with effect sizes. Grouping data by sex would afford a more powerful comparison, but few studies organized their data that way. Gender may moderate the effects of stress on immunity by virtue of the effects of sex hormones on immunity; generally, men are considered to be more biologically vulnerable (Maes, 1999), and they may be more psychosocially vulnerable (e.g.,Scanlan, Vitaliano, Ochs, Savage, & Borson, 1998).

It seems likely to us that individual differences in subjective experience also make a substantive contribution to explaining this phenomenon. Studies have convincingly demonstrated that people's cardiovascular and neuroendocrine responses to stressful experience are dependent on their appraisals of the situation and the presence of intrusive thoughts about it (Baum et al., 1993; Frankenhauser, 1975; Tomaka et al., 1997). Although the same logic should apply to people's immune responses to stressful experience, few of the studies in this area have included measures of subjective experience, and those reports were limited by methodological issues such as aggregation across heterogeneous stressors. As a consequence, measures of subjective experience were not significantly associated with immune parameters in healthy research participants, with the exception of a modest ($r = -.10$) relationship between intrusive thoughts and natural killer cell cytotoxicity. Psychological variables such as personality and emotion can give rise to individual differences in psychological and concomitant immunological

responses to stress. Optimism and coping, for example, moderated immunological responses to stressors in several studies (e.g., Barger et al., 2000; Bosch et al., 2001; Cruess et al., 2000; Segerstrom, 2001; Stowell, Kiecolt-Glaser, & Glaser, 2001).

Mechanisms of Stress Effects on the Immune System

Virtually nothing is known about the psychological pathways linking stressors with the immune system. Many theorists have argued that affect is a final common pathway for stressors (e.g., S. Cohen, Kessler, & Underwood, 1995; Miller & Cohen, 2001), yet studies have enjoyed limited success in attempting to explain people's immune responses to life experiences on the basis of their emotional states alone (Bower et al., 1998; Cole, Kemeny, Taylor, Visscher, & Fahey, 1996; Miller, Dopp, Myers, Stevens, & Fahey, 1999; Segerstrom, Taylor, Kemeny, & Fahey, 1998). Furthermore, many studies have focused on the immune effects of emotional valence (e.g., unhappy vs. happy; Futterman, Kemeny, Shapiro, & Fahey, 1994), but the immune system may be even more closely linked to emotional arousal (e.g., stimulated vs. still), especially during acute stressors (S. Cohen et al., 2000). Finally, it is possible that emotion will prove to be relatively unimportant and that other mental processes such as motivational states or cognitive appraisals will prove to be the critical psychological mechanisms linking stress and the immune system (cf. Maier, Waldstein, & Synowski, 2003).

In terms of biological mechanisms, the field is further along, but much remains to be learned. A series of studies in the mid-1990s was able to show via beta-adrenergic blockade that activation of the sympathetic nervous system was responsible for the immune system effects of acute stressors (Bachen et al., 1995; Benschop, Nieuwenhuis, et al., 1994). Apart from these findings, however, little is known about biological mechanisms, especially with regard to more enduring stressors that occur in the real world. Studies that have attempted to identify hormonal pathways linking stressors and the immune system have enjoyed limited success, perhaps because they have utilized snapshot assessments of hormones circulating in blood. Future studies can maximize their chances of identifying relevant mediators by utilizing more integrated measures of hormonal output, such as 24-hr urine collections or diurnal profiles generated through saliva collections spaced throughout the day (Baum & Grunberg, 1995; Stone et al., 2001).

Future studies could also benefit from a greater emphasis on behavior as a potential mechanism. This strategy has proven useful in studies of clinically depressed patients, in which decreased physical activity and psychomotor retardation (Cover & Irwin, 1994; Miller, Cohen, & Herbert, 1999), increased body mass (Miller, Stetler, Carney, Freedland, & Banks, 2002), disturbed sleep (Cover & Irwin, 1994; Irwin, Smith, & Gillin, 1992), and cigarette smoking (Jung & Irwin, 1999) have been shown to explain some of the immune dysregulation evident in this population. There is already preliminary evidence, for instance, that sleep loss might be responsible for some of the immune system changes that accompany stressors (Hall et al., 1998; Ironson et al., 1997).

Stress, the Immune System, and Disease

The most pressing question that future research needs to address is the extent to which stressor-induced changes in the immune system have meaningful implications for disease susceptibility in otherwise healthy humans. In the 30 years since work in the field of psychoneuroimmunology began, studies have convincingly established that stressful experiences alter features of the immune response as well as confer vulnerability to adverse medical outcomes that are either mediated by or resisted by the immune system. However, with the exception of recent work on upper respiratory infection (S. Cohen, Doyle, & Skoner, 1999), studies have not yet tied these disparate strands of work together nor determined whether immune system changes are the mechanism through which stressors increase susceptibility to

disease onset. In contrast, studies of vulnerable populations such as people with HIV have shown changes in immunity to predict disease progression (Bower et al., 1998).

To test an effect of this nature, researchers need to build clinical outcome assessments into study designs where appropriate. For example, chronic stressors reliably diminish the immune system's capacity to produce antibodies following routine influenza vaccinations (see Table 7). Yet as far as we are aware, none of these studies has tracked illness to explore whether stress-related disparities in vaccine response might be sufficient to heighten susceptibility to clinical infection with influenza. Cytokine expression represents a relatively new and promising example of an avenue for research linking stress, immune change, and disease. For example, chronic stress may elicit prolonged secretion of cortisol, to which white blood cells mount a counterregulatory response by downregulating their cortisol receptors. This downregulation, in turn, reduces the cells' capacity to respond to anti-inflammatory signals and allows cytokine-mediated inflammatory processes to flourish (Miller, Cohen, & Ritchey, 2002). Stress therefore might contribute to the course of diseases involving excessive nonspecific inflammation (e.g., multiple sclerosis, rheumatoid arthritis, coronary heart disease) and thereby increase risk for excess morbidity and mortality (Ershler & Keller, 2000;Papanicoloaou et al., 1998;Rozanski, Blumenthal, & Kaplan, 1999). Another example of the importance of cytokines to clinical pathology is in asthma and allergy, in which emerging evidence implicates excess Th2 cytokine secretion in the exacerbation of these diseases (Busse & Lemanske, 2001;Luster, 1998).

Conclusion

Sapolsky (1998) wrote,

> Stress-related disease emerges, predominantly, out of the fact that we so often activate a physiological system that has evolved for responding to acute physical emergencies, but we turn it on for months on end, worrying about mortgages, relationships, and promotions. (p. 7)

The results of this meta-analysis support this assertion in one sense: Stressors with the temporal parameters of the fight-or-flight situations faced by humans' evolutionary ancestors elicited potentially beneficial changes in the immune system. The more a stres-sor deviated from those parameters by becoming more chronic, however, the more components of the immune system were affected in a potentially detrimental way.

Further research is needed to support two other ideas elicited by this quote: the idea that subjective experience such as worry is more likely to result in stress-related immune change than objective experience and the idea that stress-related immune change results in stress-related disease. Though the results of the meta-analysis were not encouraging on the first point, many of these studies suffered from methodological limitations. We hope that these results will inform investigations that go beyond the relationship between a stressful event and an immune parameter to investigate the psychological phenomena that mediate that relationship. Finally, these results can also inform investigations into stress, immunity, and disease process. Whether the disease is characterized by natural or specific immunity, its cytokine profile, and its regulation by anti-inflammatory agents such as cortisol, may determine the disparate effects of different kinds of stressors.

Acknowledgements

Preparation of this work was supported by American Heart Association Grant 0160367Z, the National Alliance for Research on Schizophrenia and Depression, National Institute of Mental Health Grant 61531, and Michael Smith Foundation for Health Research Grant CI-SCH-58. We thank Edith Chen for her helpful comments on an earlier version of the article and Jennifer Snedeker for assistance with manuscript preparation.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Abdeljaber MH, Nair MPN, Schork MA, Schwartz SA. Depressed natural killer cell activity in schizophrenic patients. Immunological Investigations 1994;23:259–268. [PubMed: 7959960]

*Ackerman KD, Martino M, Heyman R, Moyna NM, Rabin BS. Immunologic response to acute psychological stress in MS patients and controls. Journal of Neuroimmunology 1996;68:85–94. [PubMed: 8784264]

*Ackerman KD, Martino M, Heyman R, Moyna NM, Rabin BS. Stressor-induced alteration of cytokine production in multiple sclerosis patients and controls. Psychosomatic Medicine 1998;60:484– 491. [PubMed: 9710295]

Ader R, Cohen N, Felten D. January 14). Psychoneuroimmunology: Interactions between the nervous system and the immune system. Lancet 1995;345:99–103. [PubMed: 7815892]

Ader, R., Felten, D. L., & Cohen, N. (2001). *Psychoneuroimmunology* (3rd ed.). San Diego, CA: Academic Press.

*Aloe L, Bracci-Laudiero L, Alleva E, Lambiase A, Micera A, Tirassa P. Emotional stress induced by parachute jumping enhances blood nerve growth factor levels and the distribution of nerve growth factor receptors in lymphocytes. Proceedings of the National Academy of Sciences, USA 1994;91:10440–10444.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

*Andersen BL, Farrar WB, Golden-Kreutz D, Kutz LA, MacCallum R, Courtney ME, Glaser R. Stress and immune responses after surgical treatment for regional breast cancer. Journal of the National Cancer Institute 1998;90:30–36. [PubMed: 9428780]

Andersen BL, Kiecolt-Glaser JK, Glaser R. A biobehavioral model of cancer stress and disease course. American Psychologist 1994;49:389– 404. [PubMed: 8024167]

Anstead MI, Hunt TA, Carlson SL, Burki NK. Variability of peripheral blood lymphocyte beta-2-adrenergic receptor density in humans. American Journal of Respiratory and Critical Care Medicine 1998;157:990–992. [PubMed: 9517622]

*Antoni MH, August S, LaPerriere A, Baggett HL, Klimas N, Ironson G, et al. Psychological and neuroendocrine measures related to functional immune changes in anticipation of HIV-1 serostatus notification. Psychosomatic Medicine 1990;52:496–510. [PubMed: 2247555]

*Aragona M, Muscatello MRA, Losi E, Panetta S, la Torre F, Pastura G, et al. Lymphocyte number and stress parameter modifications in untreated breast cancer patients with depressive mood and previous life stress. Journal of Experimental Therapeutics and Oncology 1996;1:354–360. [PubMed: 9414425]

*Arber N, Berliner S, Arber L, Lipshitz A, Sinai Y, Zajicek G, et al. The state of leukocyte adhesiveness/aggregation in the peripheral blood is more sensitive than the white blood cell count for the detection of acute mental stress. Journal of Psychosomatic Research 1992;36:37–46. [PubMed: 1538349]

*Arnetz BB, Brenner SO, Levi L, Petterson IL, Wasserman J, Petrini B, et al. Neuroendocrine and immunologic effects of unemployment and job insecurity. Psychotherapy and Psychosomatics 1991;55:76– 80. [PubMed: 1891571]

*Bachen EA, Manuck SB, Cohen S, Muldoon MF, Raible R, Herbert TB, Rabin BS. Adrenergic blockade ameliorates cellular immune responses to mental stress in humans. Psychosomatic Medicine 1995;57:366–372. [PubMed: 7480566]

*Bachen EA, Manuck SB, Marsland AL, Cohen S, Malkoff SB, Muldoon MF, Rabin BS. Lymphocyte subset and cellular immune responses to a brief experimental stressor. Psychosomatic Medicine 1992;54:673– 679. [PubMed: 1454961]

*Baker GHB, Byrom NA, Irani MS, Brewerton DA, Hobbs JR, Wood RJ, Nagvekar NM. March 10). Stress, cortisol, and lymphocyte subpopulations. Lancet 1984;10:574. [PubMed: 6142297]

*Baker GHB, Irani MS, Byrom NA, Nagvekar NM, Wood RJ, Hobbs JR, Brewerton DA. Stress, cortisol concentrations, and lymphocyte subpopulations. British Medical Journal 1985;290:1393. [PubMed: 3922508]

*Barger ST, Marsland AL, Bachen EA, Manuck SB. Repressive coping and blood measures of disease risk: Lipids and endocrine and immunological responses to a laboratory stressor. Journal of Applied Social Psychology 2000;30:1619–1638.

*Bartrop RW, Luckhurst E, Lazarus L, Kiloh LG, Penny R. April 16). Depressed lymphocyte function after bereavement. Lancet 1977;1:834– 836. [PubMed: 67339]

*Bauer ME, Vedhara K, Perks P, Wilcock GK, Lightman SL, Shanks N. Chronic stress in caregivers of dementia patients is associated with reduced lymphocyte sensitivity to glucocorticoids. Journal of Neuroimmunology 2000;103:84–92. [PubMed: 10674993]

Baum A, Cohen L, Hall M. Control and intrusive memories as possible determinants of chronic stress. Psychosomatic Medicine 1993;55:274–286. [PubMed: 8346335]

Baum, A., & Grunberg, N. (1995). Measurement of stress hormones. In S. Cohen, R. C. Kessler, & L. G. Underwood (Eds.), *Measuring stress: A guide for health and social scientists* (pp. 193–212). New York: Oxford University Press.

*Beck RJ, Cesario TC, Yousefi A, Enamoto H. Choral singing, performance perception, and immune system changes in salivary immunoglobulin A and cortisol. Music Perception 2000;18:87–106.

*Beem EE, Hooijkaas H, Cleiren MH, Schut HA, Garssen B, Croon MA, et al. The immunological and psychological effects of bereavement: Does grief counseling really make a difference? A pilot study. Psychiatry Research 1999;85:81–93. [PubMed: 10195319]

Ben Eliyahu S, Shakhar G, Page GG, Stefanski V, Shakhar K. Suppression of NK cell activity and of resistance to metastasis by stress: A role for adrenal catecholamines and beta-adrenoceptors. Neuroimmunomodulation 2000;8:154–164. [PubMed: 11124582]

Benjamini, E., Coico, R., & Sunshine, G. (2000). *Immunology: A short course* (4th ed.). New York: Wiley-Liss.

*Benschop RJ, Brosschot JF, Godaert GLR, de Smet MBM, Geenen R, Olff M, et al. Chronic stress affects immunologic but not cardiovascular responsiveness to acute psychological stress in humans. American Journal of Physiology 1994;266:R75–R80. [PubMed: 8304558]

Benschop RJ, Geenen R, Mills PJ, Naliboff BD, Kiecolt-Glaser JK, Herbert TB, et al. Cardiovascular and immune responses to acute psychological stress in young and old women: A meta-analysis. Psychosomatic Medicine 1998;60:290–296. [PubMed: 9625216]

*Benschop RJ, Godaert GLR, Geenen R, Brosschot JF, de Smet MBM, Olff M, et al. Relationships between cardiovascular and immunological changes in an experimental stress model. Psychological Medicine 1995;25:323–327. [PubMed: 7675920]

*Benschop RJ, Jabaaij L, Oostveen FG, Vingerhoets AJJM, Ballieux RE. The influence of psychological stress on immunoregulation of latent Epstein-Barr virus. Stress Medicine 1998;14:21–29.

*Benschop RJ, Jacobs R, Sommer B, Schürmeyer TH, Raab HR, Schmidt RE, Schedlowski M. Modulation of the immunologic response to acute stress in humans by β-blockade or benzodiazepines. Federation of American Societies for Experimental. Biology Journal 1996;10:517–524.

*Benschop RJ, Nieuwenhuis EES, Tromp EAM, Godaert GLR, Ballieux RE, van Doornen LJP. Effects of β-adrenergic blockade on immunologic and cardiovascular changes induced by mental stress. Circulation 1994;89:762–769. [PubMed: 7508828]

Benschop RJ, Rodriguez-Feuerhahn M, Schedlowski M. Catecholamine-induced leukocytosis: Early observations, current research, and future directions. Brain, Behavior, and Immunity 1996;10:77–91.

Biondi, M. (2001). Effects of stress on immune functions: An overview. In R. Ader, D. L. Felten, & N. Cohen (Eds.), *Psychoneuroimmunology* (3rd ed., pp. 189–226). San Diego, CA: Academic Press.

*Biondo M, Peronti M, Pacitti F, Pancheri P, Pacifici R, Altieri I, et al. Personality, endocrine and immune changes after eight months in healthy individuals under normal daily stress. Psychotherapy and Psychosomatics 1994;62:176–184. [PubMed: 7846261]

*Birmaher B, Rabin BS, Garcia MR, Jain U, Whiteside TL, Wissiamson DE, et al. Cellular immunity in depressed, conduct disorder, and normal adolescents: Role of adverse life events. Journal of the American Academy of Child & Adolescent Psychiatry 1994;33:671–678. [PubMed: 8056730]

*Bisselli R, Farrace S, D'Ameloi R, Fattorossi A. Influence of stress on lymphocyte subset distribution—A flow cytometric study in young student pilots. Aviation, Space, and Environmental Medicine 1993;64:116–120.

*Bongartz W, Lyncker I, Kossman KT. The influence of hypnosis on white blood cell count and urinary levels of catecholamines and vanillyl mandelic acid. Hypnos 1987;14:52– 61.

*Borella P, Bargellini A, Rovesti S, Pinelli M, Vivoli R, Solfrini V, Vivoli G. Emotional stability, anxiety, and natural killer activity under examination stress. Psychoneuroendocrinology 1999;24:613–627. [PubMed: 10399771]

Borenstein, M., & Rothstein, H. (1999). Comprehensive meta-analysis: A computer program for research synthesis [Computer software]. Englewood, NJ: Biostat.

*Boscarino JA, Chang J. Higher abnormal leukocyte and lymphocyte counts 20 years after exposure to severe stress: Research and clinical implications. Psychosomatic Medicine 1999;61:378–386. [PubMed: 10367620]

*Bosch JA, Brand HS, Ligtenberg TJM, Bermond B, Hoogstraten J, Amerongen AVN. Psychological stress as a determinant of protein levels and salivary-induced aggregation of streptococcus gordonii in human whole saliva. Psychosomatic Medicine 1996;58:374–382. [PubMed: 8827800]

*Bosch JA, de Geus EJC, Kelder A, Veerman ECI, Hoogstraten J, Amerongen AVN. Differential effects of active versus passive coping on secretory immunity. Psychophysiology 2001;38:836– 846. [PubMed: 11577907]

Bosch JA, Ring C, de Geus EJC, Veerman ECI, Amerongen AVN. Stress and secretory immunity. International Review of Neurobiology 2002;52:213–253. [PubMed: 12498106]

Boucher N, Dufeu-Duchesne T, Vicaut E, Farge D, Effros RB, Schachter F. CD28 expression in T cell aging and human longevity. Experimental Gerontology 1998;33:267–282. [PubMed: 9615924]

Bower JE, Kemeny ME, Taylor SE, Fahey JL. Cognitive processing, discovery of meaning, CD4 decline, and AIDS-related mortality among bereaved HIV-seropositive men. Journal of Consulting and Clinical Psychology 1998;66:979–986. [PubMed: 9874911]

*Boyce WT, Adams S, Tschann JM, Cohen F, Wara D, Gunnar MR. Adrenocortical and behavioral predictors of immune responses to starting school. Pediatric Research 1995;38:1009–1017. [PubMed: 8618776]

*Boyce WT, Chesterman EA, Martin N, Folkman S, Cohen F, Wara D. Immunologic changes occurring at kindergarten entry predict respiratory illnesses after the Loma Preita earthquake. Developmental and Behavioral Pediatrics 1993;14:296–303.

*Breznitz S, Ben-Zur H, Berzon Y, Weiss DW, Levitan G, Tarcic N, et al. Experimental induction and termination of acute psychological stress in human volunteers: Effects on immunological, neuroendocrine, cardiovascular, and psychological parameters. Brain, Behavior, and Immunity 1998;12:34–52.

*Bristow M, Hucklebridge FH, Clow A, Evans PD. Modulation of secretory immunoglobulin A in saliva in relation to an acute episode of stress and arousal. Journal of Psychophysiology 1997;11:248–255.

*Brosschot JF, Benschop RJ, Godaert GLR, de Smet MB, Olff M, Heijnen CJ, Ballieux RE. Effects of experimental psychological stress on distribution and function of peripheral blood cells. Psychosomatic Medicine 1992;54:394– 406. [PubMed: 1502282]

*Brosschot JF, Benschop RJ, Godaert GLR, Olff M, de Smet M, Heijnen CJ, Ballieux RE. Influence of life stress on immunological reactivity to mild psychological stress. Psychosomatic Medicine 1994;56:216–224. [PubMed: 8084967]

*Brosschot JF, Smelt D, de Smet M, Heijen CJ, Olff M, Ballieux RE, Godaert GLR. Effects of experimental psychological stress on T-lymphocytes and NK cells in man: An exploratory study. Journal of Psychophysiology 1991;5:59– 67.

*Burleson MH, Malarkey WB, Cacioppo JT, Poehlmann KM, Kiecolt-Glaser JK, Berntson GG, Glaser R. Postmenopausal hormone replacement: Effects on autonomic, neuroendocrine, and immune reactivity to brief psychological stressors. Psychosomatic Medicine 1998;60:17–25. [PubMed: 9492234]

Busse WW, Lemanske RF. Advances in immunology: Asthma. New England Journal of Medicine 2001;344:350–362. [PubMed: 11172168]

*Byrnes DM, Antoni MH, Goodkin K, Efantis-Potter J, Asthana D, Simon T, et al. Stressful events, pessimism, natural killer cell cytotoxicity, and cytotoxic/suppressor T cells in HIV positive Black women at risk for cervical cancer. Psychosomatic Medicine 1998;60:714–722. [PubMed: 9847030]

Cacioppo JT. Social neuroscience: Autonomic, neuroendocrine, and immune responses to stress. Psychophysiology 1994;31:113–128. [PubMed: 8153248]

*Cacioppo JT, Malarkey WB, Kiecolt-Glaser JK, Uchino BN, Sgoutas-Emch SA, Sheridan JF, et al. Heterogeneity in neuroendocrine and immune responses to brief psychological stressors as a function of autonomic cardiac activation. Psychosomatic Medicine 1995;57:154–164. [PubMed: 7792374]

*Cacioppo JT, Poehlmann KM, Kiecolt-Glaser JK, Malarkey WB, Burleson MH, Berntson GG, Glaser R. Cellular immune responses to acute stress in female caregivers of dementia patients and matched controls. Health Psychology 1998;17:182–189. [PubMed: 9548709]

*Caggiula AR, McAllister CG, Matthews KA, Berga SL, Owens JF, Miller AL. Psychological stress and immunological responsiveness in normally cycling, follicular-stage women. Journal of Neuroimmunology 1995;59:103–111. [PubMed: 7797611]

*Caudell KA, Gallucci BB. Neuroendocrine and immunological responses of women to stress. Western Journal of Nursing Research 1995;17:672– 692. [PubMed: 8597232]

*Chi DS, Neumann JK, Mota-Marquez M, Dubberley FA. Effects of acute stress on lymphocyte β_2-adrenoceptors in White males. Journal of Psychosomatic Research 1993;37:763–770. [PubMed: 8229907]

Chiappelli, F., Manfrini, E., Franceschi, C., Cossarizza, A., & Black, K. L. (1994). Steroid regulation of cytokines: Relevance for Th1 to Th2 shift? In E. R. de Kloet, E. C. Azmitia, & P. W. Landfield (Eds.), *Annals of the New York Academy of Sciences: Vol. 746. Brain corticosteroid receptors: Studies on the mechanism, function, and neurotoxicity of corticosteroid action* (pp. 204–215). New York: New York Academy of Sciences.

*Cohen F, Keaney KA, Zegans LS, Kemeny ME, Neuhaus JM, Stites DP. Differential immune system changes with acute and persistent stress for optimists vs. pessimists. Brain, Behavior, and Immunity 1999;13:155–174.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation for the behavioral sciences* Hillsdale, NJ: Erlbaum.

Cohen S, Doyle WJ, Skoner DP. Psychological stress, cytokine production, and severity of upper respiratory illness. Psychosomatic Medicine 1999;61:175–180. [PubMed: 10204970]

Cohen S, Frank E, Doyle WJ, Skoner DP, Rabin BS, Gwaltney JM Jr. Types of stressors that increase susceptibility to the common cold in healthy adults. Health Psychology 1998;17:214–223. [PubMed: 9619470]

*Cohen S, Hamrick N, Rodriguez MS, Feldman PJ, Rabin BS, Manuck SB. The stability of and intercorrelations among cardiovascular, immune, endocrine, and psychological reactivity. Annals of Behavioral Medicine 2000;22:171–197. [PubMed: 11211850]

Cohen S, Herbert TB. Health psychology: Psychological factors and physical disease from the perspective of human psychoneuroimmunology. Annual Review of Psychology 1996;47:113–142.

Cohen, S., Kessler, R. C., & Underwood, L. G. (1995). Strategies for measuring stress in studies of psychiatric and physical disorders. In S. Cohen, R. C. Kessler, & L. G. Underwood (Eds.), *Measuring stress: A guide for health and social scientists* (pp. 3–28). New York: Oxford University Press.

Cohen S, Miller GE, Rabin BS. Psychological stress and antibody response to immunization: A critical review of the human literature. Psychosomatic Medicine 2001;63:7–18. [PubMed: 11211068]

Cohen S, Williamson GM. Stress and infectious disease in humans. Psychological Bulletin 1991;109:5–24. [PubMed: 2006229]

Cole SW, Kemeny ME, Taylor SE, Visscher BR, Fahey JL. Accelerated course of human immunodeficiency virus infection in gay men who conceal their homosexual identity. Psychosomatic Medicine 1996;58:219–231. [PubMed: 8771621]

Cover H, Irwin M. Immunity and depression: Insomnia, retardation, and reduction of natural killer cell activity. Journal of Behavioral Medicine 1994;17:217–223. [PubMed: 8035453]

*Cruess S, Antoni M, Kilbourn K, Ironson G, Klimas N, Fletcher MA, et al. Optimism, distress, and immunologic status in HIV-infected gay men following Hurricane Andrew. International Journal of Behavioral Medicine 2000;7:160–182.

*Cruse JM, Lewis RE Jr, Bishop GR, Kliesch WF, Gaitan E, Britt R. Decreased immune reactivity and neuroendocrine alterations related to chronic stress in spinal cord injury and stroke patients. Pathobiology 1993;61:183–192. [PubMed: 8216840]

*Davidson RJ, Coe CC, Dolski I, Donzella B. Individual differences in prefrontal activation asymmetry predict natural killer cell activity at rest and in response to challenge. Brain, Behavior, and Immunity 1999;13:93–108.

*de Gucht V, Fischler B, Demanet C. Immune dysfunction associated with chronic professional stress in nurses. Psychiatry Research 1999;85:105–111. [PubMed: 10195321]

*Deinzer R, Kleineidam C, Stiller-Winkler R, Idel H, Bachg D. Prolonged reduction of salivary immunoglobulin A (sIgA) after a major academic exam. International Journal of Psychophysiology 2000;37:219–232. [PubMed: 10858568]

*Deinzer R, Schüller N. Dynamics of stress-related decrease of salivary immunoglobulin A (sIgA): Relationship to symptoms of the common cold and studying behavior. Behavioral Medicine 1998;23:161–169. [PubMed: 9494693]

*Dekaris D, Sabioncello A, Mažuran R, Rabatić S, Svoboda-Beusan I, Računica NL, TomašIć J. Multiple changes of immunologic parameters in prisoners of war. Journal of the American Medical Association 1993;270:595–599. [PubMed: 8331758]

*Delahanty DL, Dougall AL, Browning LJ, Hyman KB, Baum A. Duration of stressor and natural killer cell activity. Psychology and Health 1998;13:1121–1134.

*Delahanty DL, Dougall AL, Craig KJ, Jenkins FJ, Baum A. Chronic stress and natural killer cell activity after exposure to traumatic death. Psychosomatic Medicine 1997;59:467– 476. [PubMed: 9316178]

*Delahanty DL, Dougall AL, Hawken L, Trakowski JH, Schmitz JB, Jenkins FJ, Baum A. Time course of natural killer cell activity and lymphocyte proliferation in response to two acute stressors in healthy men. Health Psychology 1996;15:48–55. [PubMed: 8788540]

*Delahanty DL, Wang T, Maravich C, Forlenza M, Baum A. Time-of-day effects on response of natural killer cells to acute stress in men and women. Health Psychology 2000;19:39– 45. [PubMed: 10711586]

Dhabhar FS, McEwen BS. Acute stress enhances while chronic stress suppresses cell-mediated immunity in vivo: A potential role for leukocyte trafficking. Brain, Behavior, and Immunity 1997;11:286–306.

Dhabhar, F. S., & McEwen, B. S. (2001). Bidirectional effects of stress and glucocorticoid hormones on immune function: Possible explanations for paradoxical observations. In R. Ader, D. L. Felten, & N. Cohen (Eds.), *Psychoneuroimmunology* (3rd ed., pp. 301–338). San Diego, CA: Academic Press.

*Dimsdale JE, Mills P, Patterson T, Ziegler M, Dillon E. Effects of chronic stress on beta-adrenergic receptors in the homeless. Psychosomatic Medicine 1994;56:290–295. [PubMed: 7972610]

*Dobbin JP, Harth M, McCain GA, Martin RA, Cousin K. Cytokine production and lymphocyte transformation during stress. Brain, Behavior, and Immunity 1991;5:339–348.

*Dopp JM, Miller GE, Myers HF, Fahey JL. Increased natural killer-cell mobilization and cytotoxicity during marital conflict. Brain, Behavior, and Immunity 2000;14:10–26.

*Drummond PD, Hewson-Bower B. Increased psychosocial stress and decreased mucosal immunity in children with recurrent upper respiratory tract infections. Journal of Psychosomatic Research 1997;43:271–278. [PubMed: 9304553]

*Dugué B, Leppänen EA, Teppo AM, Fyhrquiist F, Gräsbeck R. Effects of psychological stress on plasma interleukins-1β and 6, C-reactive protein, tumour necrosis factor alpha, anti-diuretic hormone and serum cortisol. Scandinavian Journal of Clinical Lab Investigation 1993;53:555–561.

Dunlap WP, Cortina JM, Vaslow JB, Burke MJ. Meta-analysis of experiments with matched groups or repeated measures designs. Psychological Methods 1996;2:170–177.

*Dworsky R, Paganini-Hill A, Ducey B, Hechinger M, Parker JW. Lymphocyte immunophenotyping in an elderly population: Age, sex, and medication effects—A flow cytometry study. Mechanisms of Ageing and Development 1989;48:255–266. [PubMed: 2786599]

Effros RB, Boucher N, Porter V, Zhu X, Spaulding C, Walford RL, et al. Decline in CD28+ T cells in centenarians and in long-term T cell cultures: A possible cause for both in vivo and in vitro immunosenescence. Experimental Gerontology 1994;29:601– 609. [PubMed: 9435913]

Effros RB, Pawelec G. Replicative senescence of T cells: Does the Hayflick limit lead to immune exhaustion? Immunology Today 1997;18:450– 454. [PubMed: 9293162]

Elliot, G. R., & Eisdorfer, C. (1982). *Stress and human health: An analysis and implications of research. A study by the Institute of Medicine, National Academy of Sciences* New York: Springer Publishing.

*Endresen IM, Relling GB, Tønder O, Myking O, Walther BT, Ursin H. Brief uncontrollable stress and psychological parameters influence human plasma concentrations of IgM and complement component C3. Behavioral Medicine 1991 Winter;:167–176. [PubMed: 1793998]

Ershler WB, Keller ET. Age-associated increased interleukin-6 gene expression, late-life diseases, and frailty. Annual Review of Medicine 2000;51:245–270.

*Esterling BA, Kiecolt-Glaser JK, Bodnar JC, Glaser R. Chronic stress, social support, and persistent alterations in the natural killer cell response to cytokines in older adults. Health Psychology 1994;13:291–298. [PubMed: 7957007]

*Esterling BA, Kiecolt-Glaser JK, Glaser R. Psychosocial modulation of cytokine-induced natural killer cell activity in older adults. Psychosomatic Medicine 1996;58:264–272. [PubMed: 8771626]

*Evans DL, Leserman J, Perkins DO, Stern RA, Murphy C, Tamul K, et al. Stress-associated reductions of cytotoxic T lymphocytes and natural killer cells in asymptomatic HIV infection. American Journal of Psychiatry 1995;152:543–550. [PubMed: 7694902]

Felten SY, Felten D. Neural-immune interaction. Progress in Brain Research 1994;100:157–162. [PubMed: 7938514]

Ferguson RG, Wikby A, Maxson P, Olsson J, Johansson B. Immune parameters in a longitudinal study of a very old population of Swedish people: A comparison between survivors and nonsurvivors. Journals of Gerontology: Series A: Biological Sciences and Medical Sciences 1995;50:B378–B382.

*Fittschen B, Schultz KH, Schultz A, Raedler A, von Kerekjarto M. Changes of immunological parameters in healthy subjects under examination stress. International Journal of Neuroscience 1990;51:241–242. [PubMed: 2177737]

Frankenhauser, M. (1975). Experimental approaches to the study of catecholamines. In L. Levi (Eds.), *Emotions—Their parameters and measurement* (pp. 209–234). New York: Raven Press.

Futterman AD, Kemeny ME, Shapiro D, Fahey JL. Immunological and physiological changes associated with induced positive and negative mood. Psychosomatic Medicine 1994;56:499–511. [PubMed: 7871105]

*Geenen R, Godaert GLR, Heijnen CJ, Vianen ME, Wenting MJG, Nederhoff MGJ, Bijlsma JWJ. Experimentally induced stress in rheumatoid arthritis of recent onset: Effects on peripheral blood lymphocytes. Clinical and Experimental Rheumatology 1998;16:553–559. [PubMed: 9779302]

*Gennaro S, Fehder WP, Cnaan A, York R, Campbell DE, Gallagher PR, Douglas SD. Immune responses in mothers of term and preterm very-low-birth-weight infants. Clinical and Diagnostic Laboratory Immunology 1997;4:565–571. [PubMed: 9302206]

*Gennaro S, Fehder W, Nuamah IF, Campbell DE, Douglas SD. Caregiving to very low birthweight infants: A model of stress and immune response. Brain, Behavior, and Immunity 1997;11:201–215.

*Gerits P, DeBrabander B. Psychosocial predictors of psychological, neurochemical and immunological symptoms of acute stress among breast cancer patients. Psychiatry Research 1999;85:95–103. [PubMed: 10195320]

*Gerritsen W, Heijnen CJ, Wiegant VM, Bermond B, Frijda NH. Experimental social fear: Immunological, hormonal, and autonomic concomitants. Psychosomatic Medicine 1996;58:273–286. [PubMed: 8771627]

*Gilbert DG, Stuckard ME, Jensen RA, Detwiler FRJ, Martinko JM. Effects of exam stress on mood, cortisol, and immune functioning: Influences of neuroticism and smoker–non-smoker status. Personality and Individual Differences 1996;21:235–246.

*Glaser R, Friedman SB, Smyth J, Ader R, Bijur P, Brunell P, et al. The differential impact of training stress and final examination stress on herpesvirus latency at the United States Military Academy at West Point. Brain, Behavior, and Immunity 1999;13:240–251.

*Glaser R, Kennedy S, Lafuse WP, Bonneau RH, Speicher C, Hillhouse J, Kiecolt-Glaser JK. Psychological stress-induced modulation of interleukin 2 receptor gene expression and inter-leukin 2 production in peripheral blood leukocytes. Archives of General Psychiatry 1990;47:707–712. [PubMed: 2378541]

*Glaser R, Kiecolt-Glaser JK. Chronic stress modulates the virus-specific immune response to latent herpes simplex virus type 1. Annals of Behavioral Medicine 1997;19:78– 82. [PubMed: 9603681]

*Glaser, R., Kiecolt-Glaser, J. K., Malarkey, W. B., & Sheridan, J. F. (1998). The influence of psychological stress on the immune response to vaccines. In S. M. McCann, J. M. Lipton, E. M.

Sternberg, G. P. Chrousos, P. W. Gold, & C. C. Smith (Eds.), *Annals of the New York Academy of Sciences: Vol. 840. Neuroimmunomodulation: Molecular aspects, integrative systems, and clinical advances* (pp. 649– 655). New York: New York Academy of Sciences.

*Glaser R, Kiecolt-Glaser JK, Speicher CE, Holliday JE. Stress, loneliness, and changes in herpesvirus latency. Journal of Behavioral Medicine 1985;8:249–260. [PubMed: 3003360]

*Glaser R, Kiecolt-Glaser JK, Stout JC, Tarr KL, Speicher CE, Holliday JE. Stress-related impairments in cellular immunity. Psychiatry Research 1985;16:233–239. [PubMed: 2935896]

*Glaser R, MacCallum RC, Laskowski BF, Malarkey WB, Sheridan JF, Kiecolt-Glaser JK. Evidence for a shift in the Th-1 to Th-2 cytokine response associated with chronic stress and aging. Journal of Gerontology: Series A: Biological Sciences and Medical Sciences 2001;56:M477–M482.

*Glaser R, Mehl VS, Penn G, Speicher CE, Kiecolt-Glaser JK. Stress-associated changes in plasma immunoglobulin levels. International Journal of Psychosomatics 1986;33:41– 42. [PubMed: 3462168]

*Glaser R, Pearl DK, Kiecolt-Glaser JK, Malarkey WB. Plasma cortisol levels and reactivation of latent Epstein-Barr Virus in response to examination stress. Psychoneuroendocrinology 1994;19:765–772. [PubMed: 7991763]

*Glaser R, Pearson GR, Bonneau RH, Esterling BA, Atkinson C, Kiecolt-Glaser JK. Stress and the memory T-cell response to the Epstein-Barr virus in healthy medical students. Health Psychology 1993;12:435– 442. [PubMed: 8293726]

*Glaser R, Pearson GR, Jones JF, Hillhouse J, Kennedy S, Mao H, Kiecolt-Glaser JK. Stress-related activation of Epstein-Barr virus. Brain, Behavior, and Immunity 1991;5:219–232.

*Glaser R, Rice J, Sheridan J, Fertel R, Stout J, Speicher C, et al. Stress-related immune suppression: Health implications. Brain, Behavior, and Immunity 1987;1:7–20.

*Glaser R, Rice J, Speicher CE, Stout JC, Kiecolt-Glaser JK. Stress depresses interferon production by leukocytes concomitant with a decrease in natural killer cell activity. Behavioral Neuroscience 1996;100:675– 678. [PubMed: 2430594]

*Glaser R, Sheridan J, Malarkey WB, MacCallum RC, Kiecolt-Glaser JK. Chronic stress modulates the immune response to a pneumococcal pneumonia vaccine. Psychosomatic Medicine 2000;62:804–807. [PubMed: 11139000]

*Goebel MU, Mills PJ. Acute psychological stress and exercise and changes in peripheral leukocyte adhesion molecule expression and density. Psychosomatic Medicine 2000;62:664– 670. [PubMed: 11020096]

*Goebel MU, Mills PJ, Irwin MR, Ziegler MG. Interleukin-6 and tumor necrosis factor-α production after acute psychological stress, exercise, and infused isoproterenol: Differential effects and pathways. Psychosomatic Medicine 2000;62:591–598. [PubMed: 10949106]

*Gomez V, Zimmerman G, Froehlich WD, Knop J. Stress, control experience, acute hormonal and immune reactions. Psychologische Beiräge 1994;36:74– 81.

*González-Quijano MI, Martín M, Millán S, López-Calderón A. Lymphocyte response to mitogens: Influence of life events and personality. Neuropsychobiology 1998;38:90–96. [PubMed: 9732209]

*Goodkin K, Blaney NT, Feaster D, Fletcher MA, Baum MK, Mantero-Atienza E, et al. Active coping style is associated with natural killer cell cytotoxicity in asymptomatic HIV-1 seropositive homosexual men. Journal of Psychosomatic Research 1992;36:635– 650. [PubMed: 1403998]

*Goodkin K, Feaster DJ, Tuttle R, Blaney NT, Kumar M, Baum MK, et al. Bereavement is associated with time-dependent decrements in cellular immune function in asymptomatic human immunodeficiency virus type 1-seropositive homosexual men. Clinical Diagnosis and Lab Immunology 1996;3:109–118.

*Goodkin K, Fuchs I, Feaster D, Leeka J, Rishel DD. Life stressors and coping style are associated with immune measures in HIV-1 infection: A preliminary report. International Journal of Psychiatry in Medicine 1992;22:155–172. [PubMed: 1355469]

*Graham NMH, Bartholomeusz CA, Taboonpog N, la Brooy JT. Does anxiety reduce the secretion rate of secretory IgA in saliva? Medical Journal of Australia 1988;148:131–133. [PubMed: 3257543]

*Gruzelier J, Smith F, Nagy A, Henderson D. Cellular and humoral immunity, mood and exam stress: The influences of self-hypnosis and personality predictors. International Journal of Psychophysiology 2001;42:55–71. [PubMed: 11451479]

*Guidi L, Tricerri A, Vangeli M, Frasca D, Errani AR, Di Giovanni A, et al. Neuropeptide Y plasma levels and immunological changes during academic stress. Neuropsychobiology 1999;40:188–195. [PubMed: 10559701]

*Halim S, Kaplan HB, Pollack MS. Moderating effects of gender and vulnerability on the relationships between financial hardship, low education and immune response. Stress Medicine 2000;16:167–177.

*Hall M, Baum A, Buysse DJ, Prigerson HG, Kupfer DJ, Reynolds CF III. Sleep as a mediator of the stress-immune relationship. Psychosomatic Medicine 1998;60:48–51. [PubMed: 9492239]

*Halvorsen R, Vassend O. Effects of examination stress on some cellular immunity functions. Journal of Psychosomatic Research 1987;31:693–701. [PubMed: 2963118]

Herbert TB, Cohen S. Stress and immunity in humans: A meta-analytic review. Psychosomatic Medicine 1993;55:364–379. [PubMed: 8416086]

*Herbert TB, Cohen S, Marsland AL, Bachen EA, Rabin BS, Muldoon MF, Manuck SB. Cardiovascular reactivity and the course of immune response to an acute psychological stressor. Psychosomatic Medicine 1994;56:337–344. [PubMed: 7972616]

*Howland LC, Gotrmaker SL, Mofenson LM, Spinp C, Gardner JD, Gorski H, et al. Effects of negative life events on immune suppression in children and youth infected with human immunodeficiency virus type 1. Pediatrics 2000;106:540–546. [PubMed: 10969100]

*Inoue-Sakurai C, Maruyama S, Morimoto K. Posttraumatic stress and lifestyles are associated with natural killer cell activity in victims of the Hanshin-Awaji earthquake in Japan. Preventive Medicine 2000;31:467– 473. [PubMed: 11071826]

*Ironson G, LaPerriere A, Antoni M, O'Hearn P, Schneiderman N, Klimas N, Fletcher MA. Changes in immune and psychological measures as a function of anticipation and reaction to news of HIV-1 antibody status. Psychosomatic Medicine 1990;52:247–270. [PubMed: 1973303]

*Ironson G, Wynings C, Schneiderman N, Baum A, Rodriguez M, Greenwood D, et al. Posttraumatic stress symptoms, intrusive thoughts, loss, and immune function after Hurricane Andrew. Psychosomatic Medicine 1997;59:128–141. [PubMed: 9088048]

*Irwin M, Brown M, Patterson T, Hauger R, Mascovich A, Grant I. Neuropeptide Y and natural killer cell activity: Findings in depression and Alzheimer caregiver stress. Federation of American Societies for Experimental Biology 1991;5:3100–3107.

*Irwin M, Daniels M, Bloom ET, Smith TL, Weiner H. Life events, depressive symptoms, and immune function. American Journal of Psychiatry 1987;144:437– 441. [PubMed: 2952025]

*Irwin M, Daniels M, Bloom ET, Weiner H. Life events, depression, and natural killer cell activity. Psychopharmacology Bulletin 1986;22:1093–1096. [PubMed: 3809376]

*Irwin M, Daniels M, Risch SC, Bloom E, Weiner H. Plasma cortisol and natural killer cell activity during bereavement. Biological Psychiatry 1988;24:173–178. [PubMed: 3390497]

*Irwin M, Daniels M, Smith TL, Bloom E, Weiner H. Impaired natural killer cell activity during bereavement. Brain, Behavior, and Immunity 1987;1:98–104.

*Irwin M, Daniels M, Weiner H. Immune and neuroendocrine changes during bereavement. Psychiatric Clinics of North America 1987;10:449– 465. [PubMed: 3317313]

*Irwin M, Hauger R, Patterson TL, Semple S, Ziegler M, Grant I. Alzheimer caregiver stress: Basal natural killer cell activity, pituitary-adrenal cortical function, and sympathetic tone. Annals of Behavioral Medicine 1997;19:83–90. [PubMed: 9603682]

*Irwin M, Patterson T, Smith TL, Caldwell C, Brown SA, Gillin C, Grant I. Reduction of immune function in life stress and depression. Biological Psychiatry 1990;27:22–30. [PubMed: 2297549]

Irwin M, Smith TL, Gillin JC. Electroencephalographic sleep and natural killer cell activity in depressed patients and control subjects. Psychosomatic Medicine 1992;54:10–21. [PubMed: 1553396]

*Jabaaij L, Grosheide PM, Heijtink RA, Duivenvoorden HJ, Ballieux RE, Vingerhoets AJJM. Influence of perceived psychological stress and distress on antibody response to low dose rDNA hepatitis B vaccine. Journal of Psychosomatic Research 1993;37:361–369. [PubMed: 8510062]

*Jabaaij L, van Hattum J, Vingerhoets AJJM, Oostveen FG, Duivenvoorden HJ, Ballieux RE. Modulation of immune response to rDNA hepatitis B vaccination by psychological stress. Journal of Psychosomatic Research 1996;41:129–137. [PubMed: 8887826]

*Jacobs R, Pawlak CR, Mikeska E, Meyer-Olsen D, Martin M, Heijen CJ, et al. Systemic lupus erythematosus and rheumatoid arthritis patients differ from healthy controls in their cytokine pattern after stress exposure. Rheumatology 2001;40:868– 875. [PubMed: 11511755]

Janeway, C. A., & Travers, P. (1997). *Immunobiology: The immune system in health and disease* (3rd ed.). New York: Garland.

*Jemmott JB III, Borysenko JZ, Borysenko M, McClelland DC, Chapman R, Meyer D, Benson H. June 25). Academic stress, power motivation, and decrease in secretion rate of salivary secretory immunoglobulin A. Lancet 1983;1:1400–1402. [PubMed: 6134179]

*Jemmott JB III, Magloire K. Academic stress, social support, and secretory immunoglobulin A. Journal of Personality and Social Psychology 1988;55:803– 810. [PubMed: 3210147]

Jenkins FJ, Baum A. Stress and reactivation of latent herpes simplex virus: A fusion of behavioral medicine and molecular biology. Annals of Behavioral Medicine 1995;17:116–123.

*Jern C, Wadenvik H, Mark H, Hallgren J, Jern S. Haematological changes during acute mental stress. British Journal of Haematology 1989;71:153–156. [PubMed: 2917123]

*Johnson VC, Walker LG, Heys SD, Whiting PH, Eremin O. Can relaxation training and hypnotherapy modify the immune response to stress, and is hypnotizability relevant? Contemporary Hypnosis 1996;13:100–108.

Jung W, Irwin M. Reduction of natural killer cell cytotoxic activity in major depression: Interaction between depression and cigarette smoking. Psychosomatic Medicine 1999;61:263–270. [PubMed: 10367603]

*Kamei T, Kumano H, Iwata K, Yasushi M. Influences of long- and short-distance driving on alpha waves and natural killer cell activity. Perceptual and Motor Skills 1998;87:1419–1423. [PubMed: 10052102]

*Kamei T, Kumano H, Masumura S. Changes of immuno-regulatory cells associated with psychological stress and humor. Perceptual and Motor Skills 1997;84:1296–1298. [PubMed: 9229449]

*Kang DH, Coe CL, Karaszewski J, McCarthy DO. Relationship of social support to stress responses and immune function in healthy and asthmatic adolescents. Research in Nursing and Health 1998;21:117–128. [PubMed: 9535404]

*Kang DH, Coe CL, McCarthy DO. Academic examinations significantly impact immune responses, but not lung function, in healthy and well-managed asthmatic adolescents. Brain, Behavior, and Immunity 1996;10:164–181.

*Kang DH, Coe CL, McCarthy DO, Ershler WB. Immune responses to final exams in healthy and asthmatic adolescents. Nursing Research 1997;46:12–19. [PubMed: 9024419]

*Kang DH, Fox C. Neuroendocrine and leukocyte responses and pulmonary function to acute stressors. Annals of Behavioral Medicine 2000;22:276–285. [PubMed: 11253438]

*Kawakami N, Tanigawa T, Araki S, Nakata A, Sakurai S, Yokoyama K, Morita Y. Effects of job strain on helper-inducer (CD4+CD29+) and suppressor-inducer (CD4+CD45RA+) T cells in Japanese blue-collar workers. Psychotherapy and Psychosomatics 1997;66:192–198. [PubMed: 9259042]

*Kawamura N, Yoshiharu K, Asukai N. Suppression of cellular immunity in men with a past history of posttraumatic stress disorder. American Journal of Psychiatry 2001;158:484– 486. [PubMed: 11229994]

*Kemeny ME, Cohen F, Zegans LS, Conant MA. Psychological and immunological predictors of genital herpes recurrence. Psychosomatic Medicine 1989;51:195–208. [PubMed: 2565589]

Kemeny, M. E., Solomon, G. F., Morley, J. E., & Herbert, T. B. (1992). Psychoneuroimmunology. In C. B. Nemeroff (Ed.), *Neuroendocrinology* (pp. 563–591). Boca Raton, FL: CRC Press.

*Kessler RC, Foster C, Joseph J, Ostrow D, Wortman C, Phair J, Chmiel J. Stressful life events and symptom onset in HIV infection. American Journal of Psychiatry 1991;148:733–738. [PubMed: 1674646]

Kiecolt-Glaser JK, Cacioppo JT, Malarkey WB, Glaser R. Acute psychological stressors and short-term immune changes: What, why, for whom, and to what extent? Psychosomatic Medicine 1992;54:680– 685. [PubMed: 1454962]

*Kiecolt-Glaser JK, Dura JR, Speicher CE, Trask J, Glaser R. Spousal caregivers of dementia victims: Longitudinal changes in immunity and health. Psychosomatic Medicine 1991;53:345–362. [PubMed: 1656478]

*Kiecolt-Glaser JK, Fisher LD, Ogrocki P, Stout JC, Speicher CE, Glaser R. Marital quality, marital disruption, and immune function. Psychosomatic Medicine 1987;49:13–34. [PubMed: 3029796]

*Kiecolt-Glaser JK, Garner W, Speicher C, Penn GM, Holliday J, Glaser R. Psychosocial modifiers of immunocompetence in medical students. Psychosomatic Medicine 1994;46:7–14. [PubMed: 6701256]

Kiecolt-Glaser JK, Glaser R. Methodological issues in behavioral immunology research with humans. Brain, Behavior, and Immunity 1988;2:67–78.

*Kiecolt-Glaser JK, Glaser R, Cacioppo JT, MacCallum RC, Snydersmith M, Kim C, Malarkey WB. Marital conflict in older adults: Endocrinological and immunological correlates. Psychosomatic Medicine 1997;59:339–349. [PubMed: 9251151]

*Kiecolt-Glaser JK, Glaser R, Gravenstein S, Malarkey WB, Sheridan J. Chronic stress alters the immune response to influenza virus vaccine in older adults. Proceedings of the National Academy of Sciences, USA 1996;93:3043–3047.

*Kiecolt-Glaser JK, Glaser R, Shuttleworth EC, Dyer CS, Ogrocki P, Speicher CE. Chronic stress and immunity in family caregivers of Alzheimer's disease victims. Psychosomatic Medicine 1987;49:523–535. [PubMed: 3671639]

*Kiecolt-Glaser JK, Glaser R, Strain EC, Stout JC, Tarr KL, Holliday JE, Speicher CE. Modulation of cellular immunity in medical students. Journal of Behavioral Medicine 1986;9:5–21. [PubMed: 2939253]

*Kiecolt-Glaser JK, Kennedy S, Malkoff S, Fisher L, Speicher CE, Glaser R. Marital discord and immunity in males. Psychosomatic Medicine 1988;50:213–229. [PubMed: 2838864]

*Kiecolt-Glaser JK, Malarkey WB, Chee MA, Newton T, Cacioppo JT, Mao HY, Glaser R. Negative behavior during marital conflict is associated with immunological down-regulation. Psychosomatic Medicine 1993;55:395– 409. [PubMed: 8265740]

*Kiecolt-Glaser JK, Marucha PT, Atkinson C, Glaser R. Hypnosis as a modulator of cellular immune dysregulation during acute stress. Journal of Consulting and Clinical Psychology 2001;69:674–682. [PubMed: 11550733]

*Kiecolt-Glaser JK, Marucha PT, Malarkey WB, Mercado AM, Glaser R. November 4). Slowing of wound healing by psychological stress. Lancet 1995;346:1194–1196. [PubMed: 7475659]

Kiecolt-Glaser JK, McGuire L, Robles TF, Glaser R. Psychoneuroimmunology and *Psychosomatic Medicine*: Back to the future. Psychosomatic Medicine 2002;64:15–18. [PubMed: 11818582]

*Kubitz KA, Peavey BS, Moore BS. The effect of daily hassles of humoral immunity: An interaction moderated by locus of control. Biofeedback and Self-Regulation 1986;11:115–123. [PubMed: 3567231]

*Kugler J, Reintjes F, Tewes V, Schedlowski M. Competition stress in soccer coaches increases salivary immunoglobulin A and salivary cortisol concentration. Journal of Sports Medicine and Physical Fitness 1996;36:117–120. [PubMed: 8898518]

Kumar M, Kumar AM, Waldrop D, Antoni MH, Schneiderman N, Eisdorfer C. The HPA axis in HIV-1 infection. Journal of Acquired Immune Deficiency Syndromes 2002;31(Suppl 2):S89–S93. [PubMed: 12394788]

*Kusaka Y, Kondou H, Morimoto K. Healthy lifestyles are associated with higher natural killer cell activity. Preventive Medicine 1992;21:602– 615. [PubMed: 1279663]

*Lacey K, Zaharia MD, Griffiths J, Ravindran AV, Merali Z, Anisman H. A prospective study of neuroendocrine and immune alterations associated with the stress of an oral academic examination among graduate students. Psychoneuroendocrinology 2000;25:339–356. [PubMed: 10725611]

Landmann R. Beta-adrenergic receptors in human leukocyte sub-populations. European Journal of Clinical Investigation 1992;22:30–36. [PubMed: 1333965]

*Landmann RMA, Müller FB, Perini C, Wesp M, Erne P, Bühler FR. Changes of immunoregulatory cells induced by psychological and physical stress: Relationship to plasma catecholamines. Clinical and Experimental Immunology 1984;58:127–135. [PubMed: 6478647]

*Lane R, Ungerer J, Bernene J, Askenase P. Skin testing and erythrocyte sedimentation rate in women undergoing breast biopsy. International Journal of Psychiatry in Medicine 1983;13:37– 46. [PubMed: 6885263]

*Larson MR, Ader R, Moynihan JA. Heart rate, neuroendocrine, and immunological reactivity in response to an acute laboratory stressor. Psychosomatic Medicine 2001;63:493–501. [PubMed: 11382278]

* Lauc, G., Dabelić, S., Dumić, J., & Flögel, M. (1998). Stress and natural killer cell activity in professional soldiers. In P. Csermely (Ed.), *Annals of the New York Academy of Sciences: Vol. 851. Stress of life: From molecules to man* (pp. 526–530). New York: New York Academy of Sciences.

Laudenslager ML. The psychobiology of loss: Lessons from humans and nonhuman primates. Journal of Social Issues 1988;44:19–36.

*Laudenslager ML, Aasal R, Adler L, Berger CL, Montgomery PT, Sandberg E, et al. Elevated cytotoxicity in combat veterans with long-term posttraumatic stress disorder: Preliminary observations. Brain, Behavior, and Immunity 1998;12:74–79.

*Lerman Y, Melamed S, Shragin Y, Kushnir T, Rotgoltz Y, Shirom A, Aronson M. Association between burnout at work and leukocyte adhesiveness/aggregation. Psychosomatic Medicine 1999;61:828–833. [PubMed: 10593635]

*Leserman J, Petitto JM, Perkins DO, Folds JD, Golden RN, Evans DL. Severe stress, depressive symptoms, and changes in lymphocyte subsets in human immunodeficiency virus-infected men. Archives of General Psychiatry 1997;54:279–285. [PubMed: 9075469]

*Levy SM, Herberman RB, Simons A, Whiteside T, Lee J, McDonald R, Beadle M. Persistently low natural killer cell activity in normal adults: Immunological, hormonal and mood correlates. Natural Immunity and Cell Growth Regulation 1989;8:173–186. [PubMed: 2779599]

*Liang SW, Jemerin JM, Tshann JM, Wara DW, Boyce WT. Life events, frontal electroencephalogram laterality, and functional immune status after acute psychological stressors in adolescents. Psychosomatic Medicine 1997;59:178–186. [PubMed: 9088055]

*Linn BS, Linn MW, Klimas NG. Effects of psychophysical stress on surgical outcome. Psychosomatic Medicine 1988;50:230–244. [PubMed: 3387507]

*Linn MW, Linn BS, Jensen J. Stressful events, dysphoric mood, and immune responsiveness. Psychological Reports 1984;54:219–222. [PubMed: 6718615]

*Linn MW, Linn BS, Skyler JS, Jensen J. Stress and immune function in diabetes mellitus. Clinical Immunity and Immunopathology 1983;27:223–233.

*Lowe G, Urquhart J, Greenman J, Lowe G. Academic stress and secretory immunoglobulin A. Psychological Reports 2000;87:721–722. [PubMed: 11191373]

Luster AD. Chemokines—Chemotactic cytokines that mediate inflammation. New England Journal of Medicine 1998;338:436– 445. [PubMed: 9459648]

*Lutgendorf SK, Antoni MH, Ironson G, Klimas N, Fletcher MA, Schneiderman N. Cognitive processing style, mood, and immune function following HIV seropositivity notification. Cognitive Therapy and Research 1997;21:157–184.

*Lutgendorf SK, Reimer TT, Harvey JH, Marks G, Hong S-Y, Hillis SL, Lubaroff DM. Effects of housing relocation on immunocompetence and psychosocial functioning in older adults. Journal of Gerontology: Series A: Biological Sciences and Medical Sciences 2001;56:M97–M105.

*Lutgendorf SK, Vitaliano PP, Tripp-Reimer T, Harvey JH, Lubaroff DM. Sense of coherence moderated the relationship between life stress and natural killer cell activity in healthy older adults. Psychology and Aging 1999;14:552–563. [PubMed: 10632144]

Madden, K. S., Thyagarajan, S., & Felten, D. L. (1998). Alterations in sympathetic noradrenergic innervation in lymphoid organs with age. In S. M. McCann, J. M. Lipton, E. M. Sternberg, G. P. Chrousos, P. W. Gold, & C. C. Smith (Eds.), *Annals of the New York Academy of Sciences: Vol. 840. Neuroimmunomodulation: Molecular aspects, integrative systems, and clinical advances* (pp. 262–268). New York: New York Academy of Sciences.

Maes, M. (1999). Major depression and activation of the inflammatory response system. In R. Dantzer, E. E. Wollman, & R. Yirmiya (Eds.), *Cytokines, stress, and depression* (pp. 25– 46). New York: Kluwer.

*Maes M, Hendriks D, van Gastel A, Demedts P, Wauters A, Neels H, et al. Effects of psychological stress on serum immunoglobulin, complement and acute phase protein concentrations in normal volunteers. Psychoneuroendocrinology 1997;22:397– 409. [PubMed: 9364619]

*Maes M, Song C, Lin A, de Jongh R, van Gastel A, Kenis G, et al. The effects of psychological stress on humans: Increased production of pro-inflammatory cytokines and a Th1-like response in stress-induced anxiety. Cytokine 1998;10:313–318. [PubMed: 9617578]

*Maes M, van Bockstaele DR, van Gastel A, Song C, Schotte C, Neels H, et al. The effects of psychological stress on leukocyte subset distribution in humans: Evidence of immune activation. Neuropsychobiology 1999;39:1–9. [PubMed: 9892853]

Maier KJ, Waldstein SR, Synowski SJ. Relation of cognitive appraisal to cardiovascular reactivity, affect, and task engagement. Annals of Behavioral Medicine 2003;26:32– 41. [PubMed: 12867352]

Maier SF, Watkins LR. Cytokines for psychologists: Implications of bidirectional immune-to-brain communication for understanding behavior, mood, and cognition. Psychological Review 1998;105:83–107. [PubMed: 9450372]

Maier SF, Watkins LR, Fleshner M. Psychoneuroimmunology: The interface between behavior, brain, and immunity. American Psychologist 1994;49:1004–1017. [PubMed: 7818221]

Maisel AS, Fowler P, Rearden A, Motulsky HJ, Michel M. A new method for isolation of human lymphocyte subsets reveals differential regulation of beta-adrenergic receptors by terbutaline treatment. Clinical Pharmacology and Therapeutics 1989;46:429– 439. [PubMed: 2551559]

*Manuck SB, Cohen S, Rabin BS, Muldoon MF, Bachen EA. Individual differences in cellular immune response to stress. Psychological Science 1991;2:111–115.

*Marchesi GF, Cotani P, Santone G, DiGuiseppe S, Bartocci C, Montroni M. Psychological and immunological relationships during acute academic stress. New Trends in Experimental and Clinical Psychiatry 1989;5:5–22.

*Marshall GD Jr, Agarwall SK, Lloyd C, Cohen L, Henniger EM, Morris GJ. Cytokine dysregulation associated with exam stress in healthy medical students. Brain, Behavior, and Immunity 1998;12:297–307.

*Marsland AL, Cohen S, Rabin BS, Manuck SB. Associations between stress, trait negative affect, acute immune reactivity, and antibody response to hepatitis B injection in healthy young adults. Health Psychology 2001;20:4–11. [PubMed: 11199064]

*Marsland AL, Manuck SB, Fazzari TV, Stewart CJ, Rabin BS. Stability of individual differences in cellular immune responses to acute psychological stress. Psychosomatic Medicine 1995;57:295–298. [PubMed: 7652131]

*Marsland AL, Muldoon MF, Cohen S, Herbert TB, Bachen EA, Patterson S, et al. Lymphocyte subset redistribution during acute laboratory stress in young adults: Mediating effects on hemoconcentration. Health Psychology 1997;16:341–348. [PubMed: 9237086]

*Martin RA, Dobbin JP. Sense of humor, hassles, and immunoglobulin A: Evidence for a stress-moderating effect of humor. International Journal of Psychiatry in Medicine 1988;18:93–105. [PubMed: 3170082]

*Marucha PT, Kiecolt-Glaser JK, Favagehi M. Mucosal wound healing is impaired by examination stress. Psychosomatic Medicine 1998;60:362–365. [PubMed: 9625226]

*Matthews KA, Caggiula AR, McAllister CG, Berga SL, Owens JF, Flory JD, Miller AL. Sympathetic reactivity to acute stress and immune response in women. Psychosomatic Medicine 1995;57:564–571. [PubMed: 8600483]

McCleery JM, Bhagwagar Z, Smith KA, Goodwin GM, Cowen PJ. Modelling a loss event: Effect of imagined bereavement on the hypothalamic-pituitary-adrenal axis. Psychological Medicine 2000;30:219–223. [PubMed: 10722192]

*McClelland DC, Alexander C, Marks E. The need for power, stress, immune function, and illness among male prisoners. Journal of Abnormal Psychology 1982;91:61–70. [PubMed: 7056944]

*McClelland DC, Floor E, Davidson RJ, Saron C. Stressed power motivation, sympathetic activation, immune function, and illness. Journal of Human Stress 1980;6:11–19. [PubMed: 7391555]

*McClelland DC, Patel V, Brown D, Kelner SP Jr. Spring). The role of affiliative loss in the recruitment of helper cells among insulin-dependent diabetics. Behavioral Medicine 1991;17:5–14. [PubMed: 1827998]

*McClelland DC, Ross G, Patel V. The effects of an academic examination on salivary norepinephrine and immunoglobulin levels. Journal of Human Stress 1985;11:52–59. [PubMed: 3843108]

*McDade TW. Lifestyle incongruity, social integration, and immune function in Samoan adolescents. Social Science and Medicine 2001;53:1351–1362. [PubMed: 11676405]

*McDade TW, Stallings JF, Angold A, Costello EJ, Burleson M, Cacioppo JT, et al. Epstein-Barr virus antibodies in whole blood spots: A minimally invasive method for assessing an aspect of cell-mediated immunity. Psychosomatic Medicine 2000;62:560–567. [PubMed: 10949102]

*McDonald RD, Yagi K. A note on eosinopenia as an index of psychological stress. Psychosomatic Medicine 1960;2:149–150.

*McIntosh WA, Kaplan HB, Kubena KS, Landmann WA. Life events, social support, and immune response in elderly individuals. International Journal of Aging and Human Development 1993;37:23–36. [PubMed: 8375915]

*McKinnon W, Weisse CS, Reynolds CP, Bowles CA, Baum A. Chronic stress, leukocyte subpopulations, and humoral response to latent viruses. Health Psychology 1989;8:389– 402. [PubMed: 2555149]

*McNaughton ME, Smith LW, Patterson TL, Grant I. Stress, social support, coping resources, and immune status in elderly women. Journal of Nervous and Mental Disease 1990;178:460– 461. [PubMed: 2142212]

*Miletic ID, Schiffman SS, Miletic VD, Sattely-Miller EA. Salivary IgA secretion rate in young and elderly persons. Physiology and Behavior 1996;60:243–248. [PubMed: 8804670]

Miller GE, Cohen S. Psychological interventions and the immune system: A meta-analytic review and critique. Health Psychology 2001;20:47– 63. [PubMed: 11199066]

Miller GE, Cohen S, Herbert TB. Pathways linking major depression and immunity in ambulatory female patients. Psychosomatic Medicine 1999;61:850– 860. [PubMed: 10593638]

Miller GE, Cohen S, Ritchey AK. Chronic psychological stress and the regulation of pro-inflammatory cytokines: A glucocorticoid resistance model. Health Psychology 2002;21:531–541. [PubMed: 12433005]

*Miller GE, Dopp JM, Myers HF, Stevens SY, Fahey JL. Psychosocial predictors of natural killer cell mobilization during marital conflict. Health Psychology 1999;18:262–271. [PubMed: 10357507]

Miller GE, Stetler CA, Carney RM, Freedland KE, Banks WA. Clinical depression and inflammatory risk markers for coronary heart disease. American Journal of Cardiology 2002;90:1279–1283. [PubMed: 12480034]

*Mills PJ, Berry CC, Dimsdale JE, Ziegler MG, Nelesen RA, Kennedy BP. Lymphocyte subset redistribution in response to acute experimental stress: Effects of gender, ethnicity, hypertension, and the sympathetic nervous system. Brain, Behavior, and Immunity 1995;9:61– 69.

*Mills PJ, Dimsdale JE. The effects of acute psychological stress on cellular adhesion molecules. Journal of Psychosomatic Research 1996;41:49–53. [PubMed: 8887818]

*Mills PJ, Dimsdale JE, Nelesen RA, Dillon E. Psychologic characteristics associated with acute stressor-induced leukocyte subset redistribution. Journal of Psychosomatic Research 1996;40:417– 423. [PubMed: 8736422]

*Mills PJ, Haeri SL, Dimsdale JE. Temporal stability of acute stressor-induced changes in cellular immunity. International Journal of Psychophysiology 1995;19:287–290. [PubMed: 7558995]

*Mills PJ, Yu H, Ziegler G, Patterson T, Grant I. Vulnerable caregivers of patients with Alzheimer's disease have a deficit in circulating CD62L− T lymphocytes. Psychosomatic Medicine 1999;61:168–174. [PubMed: 10204969]

*Mills PJ, Ziegler MG, Dimsdale JE, Parry BL. Enumerative immune changes following acute stress: Effect of the menstrual cycle. Brain, Behavior, and Immunity 1995;9:190–195.

*Mills PJ, Ziegler MG, Patterson T, Dimsdale JE, Hauger R, Irwin M, Grant I. Plasma catecholamine and lymphocyte β_2-adrenergic receptor alterations in elderly Alzheimer caregivers under stress. Psychosomatic Medicine 1997;59:251–256. [PubMed: 9178336]

*Mills PJ, Zeigler MG, Rehman J, Maisel AS. Catecholamines, catecholamine receptors, cell adhesion molecules, and acute stressor-related changes in cellular immunity. Advances in Pharmacology 1998;42:587–590. [PubMed: 9327970]

*Mosnaim AD, Wolf ME, Maturana P, Mosnaim G, Puente J, Kucuk O, Gilman-Sachs A. In vitro studies of natural killer cell activity in posttraumatic stress disorder patients: Response to methionine-enkephalin challenge. Immunopharmacology 1993;25:107–116. [PubMed: 8500983]

*Moss H, Bose S, Wolters P, Brouwers P. A preliminary study of factors associated with psychological adjustment and disease course in school-age children infected with the human immunodeficiency virus. Journal of Developmental and Behavioral Pediatrics 1998;19:18–25. [PubMed: 9524301]

*Moss RB, Moss HB, Peterson R. Microstress, mood, and natural killer-cell activity. Psychosomatics 1989;30:279–283. [PubMed: 2762485]

*Moyna NM, Bodnar JD, Goldberg HR, Shurin MS, Robertson RJ, Rabin BS. Relation between aerobic fitness level and stress induced alterations in neuroendocrine and immune function. International Journal of Sports Medicine 1999;20:136–141. [PubMed: 10190776]

*Mulder CL, Antoni MH, Duivenvoorden HJ, Kauffmann RH, Goodkin K. Active confrontational coping predicts decreases clinical progression over a one-year period in HIV-infected homosexual men. Journal of Psychosomatic Research 1995;39:957–965. [PubMed: 8926605]

*Nagabhushan M, Mathews HL, Witek-Janusek L. Aberrant nuclear expression of AP-1 and NFkB in lymphocytes of women stresses by the experience of breast biopsy. Brain, Behavior, and Immunity 2001;15:78– 84.

*Nakamura H, Nagase H, Yoshida M, Ogino K. Natural killer (NK) cell activity and NK cell subsets in workers with a tendency of burnout. Journal of Psychosomatic Research 1999;46:569–578. [PubMed: 10454173]

*Nakano Y, Nakamura S, Hirata M, Harada K, Ando K, Tabuchi T, et al. Immune function and lifestyle of taxi drivers in Japan. Industrial Health 1998;36:32–39. [PubMed: 9473856]

*Nakata A, Araki S, Tanigawa T, Miki A, Sakuri S, Kawakami N, et al. Decrease of suppressor-inducer (CD4+CD45RA) T lymphocytes and increase of serum immunoglobulin G due to perceived job stress in Japanese nuclear electric power plant workers. Journal of Occupational and Environmental Medicine 2000;42:143–150. [PubMed: 10693074]

*Naliboff BD, Benton D, Solomon GF, Morley JE, Fahey JL, Bloom ET, et al. Immunological changes in young and old adults during brief laboratory stress. Psychosomatic Medicine 1991;53:121–132. [PubMed: 2031066]

*Naliboff BD, Solomon GF, Gilmore SL, Benton D, Morley JE, Fahey JL. The effects of the opiate antagonist naloxone on measures of cellular immunity during rest and brief psychological stress. Journal of Psychosomatic Research 1995;39:345–359. [PubMed: 7636778]

*Naliboff BD, Solomon GF, Gilmore SL, Fahey JL, Benton D, Pine J. Rapid changes in cellular immunity following a confrontational role-play stressor. Brain, Behavior, and Immunity 1995;9:207–219.

*Neumann JK, Chi DS. Relationship of church giving to immunological and TxPA stress response. Journal of Psychology and Theology 1999;27:43–51.

*Neumann JK, Chi DS, Flemming R II. Hematological and immunological acute mental stress responses of people who are severely and profoundly mentally retarded. Research in Developmental Disabilities 2000;21:347–353. [PubMed: 11100798]

*Neumann JK, Quillen JH, Chi DS, Quillen JH. Physiological stress response and psychological differences as a possible function of perceived paternal religious value similarity and church attendance. Journal of Psychology and Christianity 1998;17:233–247.

*Ockenfels MC, Stierle G, Stone AA, Hellhammer D. The effect of academic examinations on herpes simplex virus-1 (HSV-1) and adenovirus latency. Psychologische Beiräge 1994;36:61– 68.

*Ohira H, Watanabe Y, Kobayashi K, Kawai M. The type A behavior pattern and immune reactivity to brief stress: Change of volume of secretory immunoglobulin A in saliva. Perceptual and Motor Skills 1999;89:423– 430. [PubMed: 10597578]

O'Leary A. Stress, emotion, and human immune function. Psychological Bulletin 1990;108:363–382. [PubMed: 2270233]

*Olff M, Brosschot JF, Godaert G, Benschop RJ, Ballieux RE, Heijnen CJ, et al. Modulatory effects of defense and coping on stress-induced changes in endocrine and immune parameters. International Journal of Behavioral Medicine 1995;2:85–103. [PubMed: 16250779]

*Paik IH, Toh KY, Lee C, Kim JJ, Lee SJ. Psychological stress may induce increased humoral and decreased cellular immunity. Behavioral Medicine 2000;26:139–141. [PubMed: 11209594]

Papanicoloaou DA, Wilder RL, Manolagas SC, Chrousos GP. The pathophysiologic roles of interleukin-6 in human disease. Annals of Internal Medicine 1998;128:127–137. [PubMed: 9441573]

*Pariante CM, Carpiniello B, Orrù MG, Sitzia R, Piras A, Farci AMG, et al. Chronic caregiving stress alters peripheral blood immune parameters: The role of age and severity of stress. Psychotherapy and Psychosomatics 1997;66:199–207. [PubMed: 9259043]

*Patterson TL, Semple SJ, Temoshok LR, Atkinson JH, Mc-Dutchan JA, Straits-Tröster K, et al. Stress and depressive symptoms prospectively predict immune change among HIV-seropositive men. Psychiatry 1995;58:299–312. [PubMed: 8746489]

*Pawlak CR, Jacobs R, Mikeska E, Ochsmann S, Lombardi MS, Kavelaars A, et al. Patients with systemic lupus erythematosus differ from healthy controls in their immunological response to acute psychological stress. Brain, Behavior, and Immunity 1999;13:287–302.

*Pawlak CR, Mikeska E, Jacobs R, Ochsmann J, Wollenhaupt J, Kavelaars A, et al. Alterations in immunological and endocrine function after acute psychological stress in patients with systemic lupus erythematosus and healthy controls. Psychologische Beiräge 2000;42:24–32.

*Pehlivanođlu B, Balkanci ZD, Ridvanađaođlu AY, Durmazlar N, Öztürk G, Erbaş D, Okur H. Impact of stress, gender, and menstrual cycle on immune system: Possible role of nitric oxide. Archives of Physiology and Biochemistry 2001;109:383–387. [PubMed: 11935378]

*Perry S, Fishman B, Jacobsberg L, Frances A. Relationships over 1 year between lymphocyte subsets and psychosocial variables among adults with infection by human immunodeficiency virus. Archives of General Psychiatry 1992;49:396– 401. [PubMed: 1586275]

*Peters ML, Godaert GLR, Ballieux RE, Brosschot JF, Sweep FCGJ, Swinkels LMJW, et al. Immune responses to experimental stress: Effects of mental effort and uncontrollability. Psychosomatic Medicine 1999;61:513–524. [PubMed: 10443760]

*Petrey LJ, Weems LB II, Livingstone JN II. Relationship of stress, distress, and the immunologic response to a recombinant hepatitis B vaccine. Journal of Family Practice 1991;32:481– 486. [PubMed: 1827147]

*Pettingale KW, Hussein M, Inayat Q, Tee DEH. Changes in immune status following conjugal bereavement. Stress Medicine 1994;10:145–150.

*Pike JL, Smith TL, Hauger RL, Nicassio PM, Patterson TL, McClintick J, et al. Chronic life stress alters sympathetic, neuroendocrine, and immune responsivity to an acute psychological stressor in humans. Psychosomatic Medicine 1997;59:447– 457. [PubMed: 9251165]

Quan N, Avitsur R, Stark JL, He L, Shah M, Caligiuri M, et al. Social stress increases the susceptibility to endotoxic shock. Journal of Neuroimmunology 2001;115:36– 45. [PubMed: 11282152]

Rabin, B. S. (1999). *Stress, immune function, and health: The connection* New York: Wiley.

*Rabkin JG, Williams JBW, Remien RH, Goetz R, Kertzner R, Gorman JM. Depression, distress, lymphocyte subsets, and human immunodeficiency virus symptoms on two occasions in HIV-positive homosexual men. Archives of General Psychiatry 1991;48:111–119. [PubMed: 1671196]

Raudenbush, S. W. (1994). Random effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301–322). New York: Russell Sage Foundation.

*Ravindran AV, Griffiths J, Merali Z, Anisman H. Primary dysthymia: A study of several psychosocial, endocrine and immune correlates. Journal of Affective Disorders 1996;40:73– 84. [PubMed: 8882917]

*Redwine LS, Altemus M, Leong YM, Carter CS. Lymphocyte responses to stress in postpartum women: Relationship to vagal tone. Psychoneuroendocrinology 2001;26:241–251. [PubMed: 11166487]

*Ring C, Harrison LK, Winzer A, Carroll D, Drayson M, Kendall M. Secretory immunoglobulin A and cardiovascular reactions to mental arithmetic, cold pressor, and exercise: Effects of alpha-adrenergic blockade. Psychophysiology 2000;37:634– 643. [PubMed: 11037040]

*Rohleder N, Schommer NC, Hellhammer DH, Engel R, Kirschbaum C. Sex differences in glucocorticoid sensitivity of proinflammatory cytokine production after psychosocial stress. Psychosomatic Medicine 2001;63:966–972. [PubMed: 11719636]

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* (Rev. ed.) Newbury Park, CA: Sage.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). New York: Russell Sage Foundation.

Rozanski A, Blumenthal JA, Kaplan JR. Impact of psychological factors on the pathogenesis of cardiovascular disease and implications for therapy. Circulation 1999;99:2192–2217. [PubMed: 10217662]

*Sabioncello A, Kocijan-Hercigonja D, Rabatić S, Tomašić J, Jeren T, Matijević L, et al. Immune, endocrine, and psychological responses in civilians displaced by war. Psychosomatic Medicine 2000;62:502–508. [PubMed: 10949095]

Sapolsky, R. M. (1998). *Why zebras don't get ulcers: An updated guide to stress, stress-related disease, and coping* New York: Freeman.

*Sauer J, Polack E, Wikindki S, Holsboer F, Stalla GK, Arzt E. The glucocorticoid sensitivity of lymphocytes changes according to the activity of the hypothalamus-pituitary-adrenocortical system. Psychoneuroendocrinology 1995;20:269–280. [PubMed: 7777655]

*Scanlan JM, Vitaliano PP, Ochs H, Savage MV, Borson S. CD4 and CD8 counts are associated with interactions of gender and psychosocial stress. Psychosomatic Medicine 1998;60:644– 653. [PubMed: 9773772]

*Schaubroeck J, Jones JR, Xie JL. Individual differences in utilizing control to cope with job demands: Effects on susceptibility to infectious disease. Journal of Applied Social Psychology 2001;86:265–278.

*Schedlowski M, Jacobs R, Alker J, Pröhl F, Stratmann G, Richter S, et al. Psychophysiological, neuroendocrine and cellular immune reactions under psychological stress. Neuropsychobiology 1993;28:87–90. [PubMed: 8255416]

*Schedlowski M, Jacobs R, Stratmann G, Richter S, Hadicke A, Tewes U, et al. Changes of natural killer cells during acute psychological stress. Journal of Clinical Immunology 1993;13:119–126. [PubMed: 8320310]

*Schlesinger M, Yodfat Y. Effect of psychosocial stress on natural killer cell activity. Cancer Detection and Prevention 1988;12:9–14. [PubMed: 3263203]

*Schlesinger M, Yodfat Y. The impact of stressful life events on natural killer cells. Stress Medicine 1991;7:53– 60.

*Schmid-Ott G, Jacobs R, Jäger B, Klages S, Wolf J, Werfel T, et al. Stress-induced endocrine and immunological changes in psoriasis patients and healthy controls. Psychotherapy and Psychosomatics 1998;67:37– 42. [PubMed: 9491439]

*Schmid-Ott G, Jaeger B, Adamek C, Koch H, Lamprecht F, Kapp A, Werfel T. Levels of circulating CD8+ T lymphocytes, natural killer cells, and eosinophils increase upon acute psychosocial stress in patients with atopic dermatitis. Journal of Allergy and Clinical Immunology 2001;107:171–177. [PubMed: 11150008]

*Segerstrom SC. Optimism, goal conflict, and stressor-related immune change. Journal of Behavioral Medicine 2001;24:441– 467. [PubMed: 11702359]

*Segerstrom SC, Taylor SE, Kemeny ME, Fahey JL. Optimism is associated with mood, coping, and immune change in response to stress. Journal of Personality and Social Psychology 1998;74:1646–1655. [PubMed: 9654763]

Selye, H. (1975). *The stress of life* New York: McGraw-Hill.

*Sgoutas-Emch SA, Cacioppo JT, Uchino BN, Malarkey W, Pearl D, Kiecolt-Glaser JK, Glaser R. The effects of an acute psychological stressor on cardiovascular, endocrine, and cellular immune response: A prospective study of individuals high and low in heart rate reactivity. Psychophysiology 1994;31:264–271. [PubMed: 8008790]

Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 261–281). New York: Russell Sage Foundation.

Shavit Y, Terman GW, Martin FC, Lewis JW, Liebeskind JC, Gale RP. Stress, opioid peptides, the immune system, and cancer. Journal of Immunology 1985;135(Suppl 2):834s– 837s.

*Shea J, Clover K, Burton R. Relationships between measures of acute and chronic stress and cellular immunity. Medical Science Research 1991;19:221–222.

Sheridan, J. F., Dobbs, C. M., Jung, J., Chu, X., Konstantinos, A., Padgett, D. A., & Glaser, R. (1998). Stress-induced neuroendocrine modulation of viral pathogenesis and immunity. In S. M. McCann, J. M. Lipton, E. M. Sternberg, G. P. Chrousos, P. W. Gold, & C. C. Smith (Eds.), *Annals of the New York Academy of Sciences: Vol. 840. Neuroimmunomodulation: Molecular aspects, integrative systems, and clinical advances* (pp. 803– 808). New York: New York Academy of Sciences.

*Sieber WJ, Rodin J, Larson L, Ortega S, Cummings N, Levy S, et al. Modulation of human natural killer cell activity by exposure to uncontrollable stress. Brain, Behavior, and Immunity 1992;6:141–156.

*Söderfeldt M, Söderfeldt B, Ohlson CG, Theorell T, Jones I. The impact of sense of coherence and high-demand/low-control job environment on self-reported health, burnout and psychophysiological stress indicators. Work & Stress 2000;14:1–15.

*Solomon GF, Segerstrom SC, Grohr P, Kemeny M, Fahey J. Shaking up immunity: Psychological and immunologic changes after a natural disaster. Psychosomatic Medicine 1997;59:114–127. [PubMed: 9088047]

*Song C, Kenis G, van Gastel A, Bosmans E, Lin A, de Jong R, et al. Influence of psychological stress on immuneinflammatory variables in normal humans. Part II. Altered serum concentrations of natural anti-inflammatory agents and soluble membrane antigens of monocytes and T lymphocytes. Psychiatry Research 1999;85:293–303. [PubMed: 10333381]

*Spangler G. Psychological and physiological responses during an exam and their relation to personality characteristics. Psychoneuroendocrinology 1997;22:423– 441. [PubMed: 9364621]

*Spivak B, Shohat B, Mester R, Avraham S, Gil-Ad I, Bleich A, et al. Elevated levels of serum interleukin-1β in combat-related posttraumatic stress disorder. Biological Psychiatry 1997;42:345–348. [PubMed: 9276074]

*Spratt ML, Denney DR. Immune variables, depression, and plasma cortisol over time in suddenly bereaved parents. Journal of Neuropsychiatry and Clinical Neurosciences 1991;3:299–306. [PubMed: 1821246]

Stone AA, Schwartz J, Smyth J, Kirschbaum C, Cohen S, Hell-hammer D, Grossman S. Individual differences in the diurnal cycle of salivary free cortisol: A replication of flattened cycles for some individuals. Psychoneuroendocrinology 2001;26:295–306. [PubMed: 11166492]

*Stone AA, Valdimarsdottir HB, Katkin ES, Burns J, Cox DS, Lee S, et al. Effects of mental stressors on mitogen-induced lymphocyte responses in the laboratory. Psychology and Health 1993;8:269–284.

*Stowell JR, Kiecolt-Glaser JK, Glaser R. Perceived stress and cellular immunity: When coping counts. Journal of Behavioral Medicine 2001;24:323–339. [PubMed: 11523331]

*Theorell T, Orth-Gomér K, Eneroth P. Slow-reacting immunoglobulin in relation to social support and changes in job strain: A preliminary note. Psychosomatic Medicine 1990;52:511–516. [PubMed: 2247556]

*Thomason B, Jones G, McClure J, Brantley P. Psychosocial co-factors in HIV illness: An empirically-based model. Psychology and Health 1996;11:385–393.

*Thornton S, Troop M, Burgess AP, Button J, Goodall R, Flynn R, et al. The relationship of psychological variables and disease progression among long-term HIV-infected men. International Journal of STD & AIDS 2000;11:734–742. [PubMed: 11089788]

*Tjemsland L, Søreide JA, Matre R, Malt UF. Preoperative psychological variables predict immunological status in patients with operable breast cancer. Psycho-Oncology 1997;6:311–320. [PubMed: 9451750]

Tomaka J, Blascovich J, Kibler J, Ernst JM. Cognitive and physiological antecedents of threat and challenge appraisal. Journal of Personality and Social Psychology 1997;73:63–72. [PubMed: 9216079]

*Tsopanakis C, Tsopanakis A. Stress hormonal factors, fatigue, and antioxidant responses to prolonged speed driving. Pharmacology, Biochemistry and Behavior 1998;60:747–751.

*Uchakin PN, Tobin B, Cubbage M, Marshall G Jr, Sams C. Immune responsiveness following academic stress in first-year medical students. Journal of Interferon and Cytokine Research 2001;21:687–694. [PubMed: 11576463]

*Uchino BN, Cacioppo JT, Malarkey W, Glaser R. Individual differences in cardiac sympathetic control predict endocrine and immune responses to acute psychological stress. Journal of Personality and Social Psychology 1995;69:736–743. [PubMed: 7473028]

*Udelman DL. Stress and immunity. Psychotherapy and Psychosomatics 1982;37:176–184. [PubMed: 6983697]

*Værnes RJ, Myhre G, Aas H, Homnes T, Hansen I, Tønder O. Relationships between stress, psychological factors, health, and immune levels among military aviators. Work & Stress 1991;5:5–16.

*Van der Pompe G, Antoni MH, Duivenvoorden HJ, Heijnen CJ. Relations of plasma ACTH and cortisol levels with the distribution and function of peripheral blood cells in response to a behavioral challenge in breast cancer: An empirical exploration by means of statistical modeling. International Journal of Behavioral Medicine 1997;4:145–169. [PubMed: 16250736]

*Van der Pompe G, Antoni MH, Visser A, Heijnen CJ. Effect of mild acute stress on immune cell distribution and natural killer cell activity in breast cancer patients. Biological Psychology 1998;48:21–35. [PubMed: 9676357]

*Van der Voort CR, Heijnen CJ, Wulffraat N, Kuis W, Kavelaars A. Stress induces increases in IL-6 production by leucocytes of patients with the chronic inflammatory disease juvenile rheumatoid arthritis: A putative role for α_1-adrenergic receptors. Journal of Neuroimmunology 2000;110:223–229. [PubMed: 11024553]

*Van Rood Y, Goulmy E, Blokland E, Pool J, van Rood J, van Houwelingen H. Stress related changes in immunological and psychological variables induced by the preparation and defense of a PhD-thesis. Psychology and Health 1995;10:229–244.

*Vassend O, Halvorsen R. Personality, examination stress, and serum concentrations of immunoglobulins. Scandinavian Journal of Psychology 1987;28:233–241. [PubMed: 3441770]

*Vedhara K, Cox NKM, Wilcock GK, Perks P, Hunt M, Anderson S, et al. February 20). Chronic stress in elderly carers of dementia patients and antibody response to influenza vaccination. Lancet 1999;353:627– 631. [PubMed: 10030328]

*Vedhara K, Nott K. The assessment of the emotional and immunological consequences of examination stress. Journal of Behavioral Medicine 1996;19:467– 478. [PubMed: 8904729]

*Vialettes B, Ozanon JP, Kaplansky S, Farnarier C, Sauvaget E, Lassmann-Vague V, et al. Stress antecedents and immune status in recently diagnosed type 1 (insulin dependent) diabetes mellitus. Diabete & Metabolisme 1989;15:45–50.

*Vitaliano PP, Scanlan JM, Ochs HD, Syrjala K, Siegler IC, Snyder EA. Psychosocial stress moderates the relationship of cancer history with natural killer cell activity. Annals of Behavioral Medicine 1998;20:199–208. [PubMed: 9989327]

*Wadee AA, Kuschke RH, Kometz S, Berk M. Personality factors, stress, and immunity. Stress and Health 2001;17:25– 40.

*Wang T, Delahanty DL, Dougall AL, Baum A. Responses of natural killer cell activity to acute laboratory stressors in healthy men at different times of day. Health Psychology 1998;17:428– 435. [PubMed: 9776001]

*Watson PB, Muller HK, Jones IH, Bradley AJ. Cell-mediated immunity in combat veterans with posttraumatic stress disorder. Medical Journal of Australia 1983;159:513–515. [PubMed: 8412949]

Wayne SJ, Rhyne RL, Garry PJ, Goodwin JS. Cell-mediated immunity as a predictor of morbidity and mortality in subjects over 60. Journals of Gerontology: Series A: Biological Sciences and Medical Sciences 1990;45:M45–M48.

*Weiss DW, Hirt R, Tarcic N, Berzon Y, Ben-Zur H, Breznitz S, et al. Studies in psychoneuroimmunology: Psychological, immunological, and neuroendocrinological parameters in Israeli civilians during and after a period of Scud missile attacks. Behavioral Medicine 1996;22:5–14. [PubMed: 8805956]

*Weisse CS, Pato CN, McAllister CG, Littman R, Breier A, Paul SM, Baum A. Differential effects of controllable and uncontrollable acute stress on lymphocyte proliferation and leukocyte percentages in humans. Brain, Behavior, and Immunity 1990;4:339–351.

*Whitehouse WG, Dinges DF, Orne EC, Keller SE, Bates BL, Bauer NK, et al. Psychosocial and immune effects of self-hypnosis for stress management throughout the first semester of medical school. Psychosomatic Medicine 1996;58:249–263. [PubMed: 8771625]

Wikby A, Johansson B, Ferguson F, Olsson J. Age-related changes in immune parameters in a very old population of Swedish people: A longitudinal study. Experimental Gerontology 1994;29:531–541. [PubMed: 7828662]

*Wilcox S, King AC, Vitaliano PP, Brassington GS. Anger expression and natural killer cell activity in family caregivers participating in a physical activity trial. Journal of Health Psychology 2000;5:431– 440.

*Willemsen G, Ring C, Carroll D, Evans P, Clow A, Hucklebridge F. Secretory immunoglobulin A and cardiovascular reactions to mental arithmetic and cold pressor. Psychophysiology 1998;35:252–259. [PubMed: 9564745]

Williams, N. A., & Leaper, D. J. (1998). Infection. In D. J. Leaper & K. G. Harding (Eds.), *Wounds: Biology and management* (pp. 71– 87). New York: Oxford University Press.

*Wilson SN, van der Kolk B, Burbridge J, Fisler R, Kradin R. Phenotype of blood lymphocytes in PTSD suggests chronic immune activation. Psychosomatics 1999;40:222–225. [PubMed: 10341534]

*Winzer A, Ring C, Carroll D, Willemsen G, Drayson M, Kendall M. Secretory immunoglobulin A and cardiovascular reactions to mental arithmetic, cold pressor, and exercise: Effects of beta-adrenergic blockade. Psychophysiology 1999;36:591– 601. [PubMed: 10442027]

*Wolf TM, Cole B, Fahrion S, Norris P, Coyne L. Age and sex modulate effects of stress on the immune system: A multivariate analysis. International Journal of Neuroscience 1994;79:121–132. [PubMed: 7744547]

*Workman EA, La Via MF. T-lymphocyte polyclonal proliferation: Effects of stress and stress response style on medical students taking national board examinations. Clinical Immunology and Immunopathology 1987;43:308–313. [PubMed: 3495378]

Yehuda R. Biology of posttraumatic stress disorder. Journal of Clinical Psychiatry 2001;62:41– 46. [PubMed: 11495096]

Yehuda R, McFarlane AC, Shalev AY. Predicting the development of posttraumatic stress disorder from the acute response to a traumatic event. Biological Psychiatry 1998;44:1305–1313. [PubMed: 9861473]

*Zakowski SG. The effects of stressor predictability on lymphocyte proliferation in humans. Psychology and Health 1995;10:409– 425.

*Zakowski SG, Cohen L, Hall MH, Wollman K, Baum A. Differential effects of active and passive laboratory stressors on immune function in healthy men. International Journal of Behavioral Medicine 1994;1:163–184. [PubMed: 16250811]

*Zakowski SG, McAllister CG, Deal M, Baum A. Stress, reactivity, and immune function in healthy men. Health Psychology 1992;11:223–232. [PubMed: 1396490]

*Zautra AJ, Okun MA, Robinson SE, Lee D, Roth SH, Emmanual J. Life stress and lymphocyte alterations among patients with rheumatoid arthritis. Health Psychology 1989;8:1–14. [PubMed: 2565227]

*Zeier H, Brauchli P, Joller-Jemelka HI. Effects of work demands on immunoglobulin A and cortisol in air traffic controllers. Biological Psychology 1996;42:413– 423. [PubMed: 8652756]

*Zisook S, Shuchter SR, Irwin M, Darko DF, Sledge P, Resivsky K. Bereavement, depression, and immune function. Psychiatry Research 1994;52:1–10. [PubMed: 8047615]

Zorrilla EP, Luborsky L, McKay JR, Rosenthal R, Houldin A, Tax A, et al. The relationship of depression and stressors to immunological assays: A meta-analytic review. Brain, Behavior, and Immunity 2001;15:199–226.

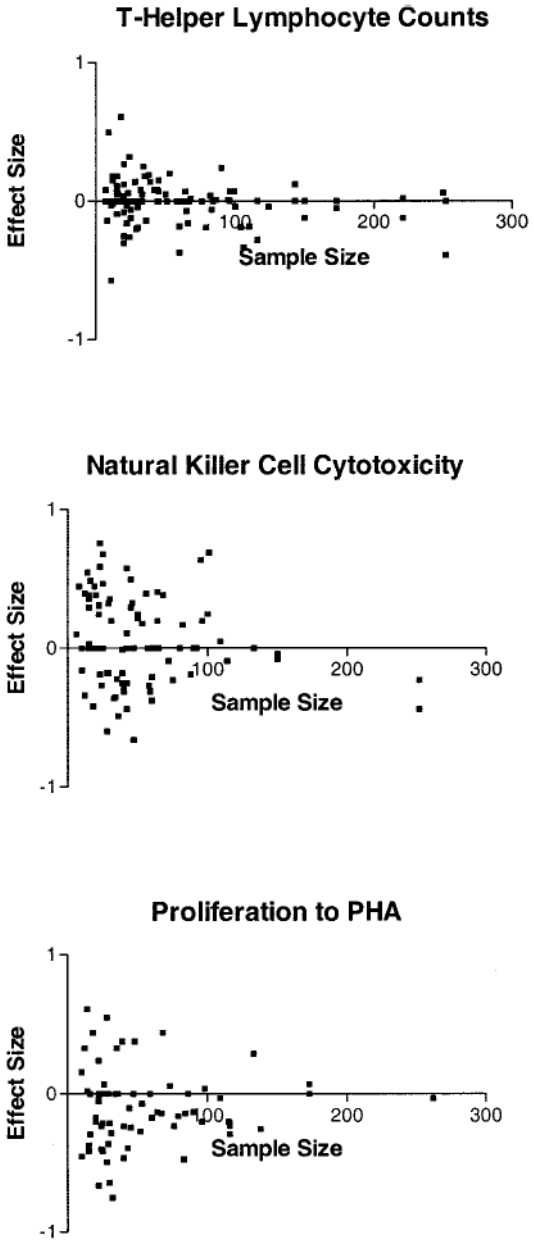


Figure 1.
Funnel plots depicting relationship between effect size and sample size. PHA = phytohemagglutinin.

Table 1

Immune Parameters Reported and Critical Characteristics

Parameter	Arm of immune system	Function	Cell surface marker
Cell			
Leukocytes	Natural	All white cells	
Granulocytes	Natural	Inflammation	
Neutrophils	Natural	Inflammation, phagocytosis	
Eosinophils	Natural	Inflammation	
Monocytes/macrophages	Natural	Inflammation, phagocytosis	
Lymphocytes	Specific	All lymphocytes	CD2
T lymphocytes	Specific	Cellular immunity	CD3, CD45RA (naive)
T-helper lymphocytes	Specific	Cellular (Th1) or humoral (Th2) immunity	CD4
T-cytotoxic lymphocytes	Specific	Cellular (Th1) immunity	CD8
B lymphocytes	Specific	Humoral (Th2) immunity	CD19, CD20
Activated B lymphocytes	Specific	Humoral (Th2) immunity	CD23, CD30
Natural killer cells	Natural	Cellular (Th1) immunity	CD16, CD56, CD57
Immunoglobulin			
IgA, IgG, IgM	Specific	Humoral (Th2) immunity	
Anti-EBV IgG	Specific	Index of EBV replication/activation	
Anti-HSV IgG	Specific	Index of HSV replication/activation	
Anti-influenza IgG postimmunization	Specific	Humoral (Th2) immunity	
Cytokine			
Interleukin-1β	Natural	Inflammation, T cell activation	
Interleukin-2	Specific	T cell activation (Th1)	
Interleukin-4	Specific	B cell activation, antibody production (Th2)	
Interleukin-6	Natural	Inflammation	
Interleukin-10	Specific	Inhibits T cell activation (Th2)	
Interferon-γ	Natural and specific	Macrophage, natural killer cell, and T cell activation (Th1)	
Tumor necrosis factor-α	Natural	Inflammation	
Complement	Natural	Increases effectiveness of natural immunity	C3
Functional assay			
Neutrophil superoxide release	Natural	Inflammation	
Natural killer cell cytotoxicity	Natural	Cellular (Th1) immunity	
Proliferation to ConA	Specific	Cellular (Th1) immunity (T cell proliferation)	
Proliferation to PHA	Specific	Cellular (Th1) immunity (T cell proliferation)	
Proliferation to PWM	Specific	Cellular (Th1) and humoral (Th2) immunity (T and B cell proliferation)	

Note. Th1 = cells that direct a response to intracellular pathogens; Th2 = cells that direct a response to extracellular pathogens; IgA = immunoglobulin A; IgG = immunoglobulin G; IgM = immunoglobulin M; EBV = Epstein-Barr virus; HSV = herpes simplex virus; ConA = concanavalin A; PHA = phytohemagglutinin; PWM = pokeweed mitogen.

Table 2

Studies Used in the Meta-Analysis by Type of Stressor

Acute time-limited	Brief naturalistic	Event sequence	Chronic	Distant	Life event	Stress appraisal
Ackerman et al., 1996, 1998	Baker et al., 1984, 1985	Antoni et al., 1990	Bauer et al., 2000	Boscarino & Chang, 1999	Abdeljaber et al., 1994	Andersen et al., 1998
Aloe et al., 1994	Bisselli et al., 1993	Aragona et al., 1996	Dekaris et al., 1993	Inoue-Sakurai et al., 2000	Benschop, Jabaaij, et al., 1998	de Gucht et al., 1999
Arber et al., 1992	Borella et al., 1999	Arnetz et al., 1991	Dimsdale et al., 1994	Laudenslager et al., 1998	Biondo et al., 1994	Halim et al., 2000
Bachen et al., 1992, 1995	Bosch et al., 1996	Bartrop et al., 1977	Drummond & Hewson-Bower 1997	Mosnaim et al., 1993	Birmaher et al., 1994	Hall et al., 1998
Barger et al., 2000	Boyce et al., 1993, 1995	Beem et al., 1999	Esterling et al., 1994, 1996	Spivak et al., 1997	Byrnes et al., 1998	Ironson et al., 1997
Beck et al., 2000	Davidson et al., 1999	Cruess et al., 2000	Gennaro, Fehder, Cnaan, et al., 1997	Watson et al., 1983	F. Cohen et al., 1999	Kawakami et al., 1997
Benschop, Brosschot, et al., 1994	Deinzer & Schüller, 1998	Delahanty et al., 1997	Gennaro, Fehder, Nuamah, et al., 1997	Wilson et al., 1999	Evans et al., 1995	Kawamura et al., 2001
Benschop et al., 1995	Deinzer et al., 2000	Dworsky et al., 1989	Glaser & Kiecolt-Glaser, 1997		Gomez et al., 1994	Kusaka et al., 1992
Benschop, Jacobs, et al., 1996	Dobbin et al., 1991	Goodkin et al., 1996	Glaser et al., 1998, 2000, 2001		González-Quijano et al., 1998	Lerman et al., 1999
Benschop, Nieuwenhuis, et al., 1994	Fittschen et al., 1990	Ironson et al., 1990, 1997	Irwin et al., 1991, 1997		Goodkin, Blaney, et al., 1992	Maes et al., 1999
Bongartz et al., 1987	Gilbert et al., 1996	Irwin et al., 1986, 1988	Kiecolt-Glaser et al., 1991, 1995, 1996		Goodkin, Fuchs, et al., 1992	Marsland et al., 2001
Bosch et al., 2001	Glaser, Kiecolt-Glaser, Speicher, & Holliday, 1985	Irwin, Daniels, Smith, et al., 1987	Kiecolt-Glaser, Glaser, et al., 1987		Graham et al., 1988	McClelland et al., 1982
Breznitz et al., 1998	Glaser, Kiecolt-Glaser, Stout, et al., 1985	Irwin, Daniels, & Weiner, 1987	Lauc et al., 1998		Howland et al., 2000	McDade, 2001
Bristow et al., 1997	Glaser et al., 1986, 1987, 1990, 1991, 1993, 1994, 1996, 1999	Kiecolt-Glaser, Fisher, et al., 1987	Lutgendorf et al., 1999		Irwin, Daniels, Bloom, et al., 1987	Nakamura et al., 1999
Brosschot et al., 1991, 1992, 1994	Gruzelier et al., 2001	Kiecolt-Glaser et al., 1988	McKinnon et al., 1989		Irwin et al., 1990	Nakata et al., 2000
Burleson et al., 1998	Guidi et al., 1999	Lane et al., 1983	Mills et al., 1997, 1999		Jabaaij et al., 1993, 1996	Scanlan et al., 1998
Cacioppo et al., 1995, 1998	Halvorsen & Vassend, 1987	Lutgendorf et al., 1997, 2001	Nakano et al., 1998		Kemeny et al., 1989	Schaubroeck et al., 2001
Caggiula et al., 1995	Jemmott & Magloire, 1988	McClelland et al., 1991	Pariante et al., 1997		Kessler et al., 1991	Söderfeldt et al., 2000
Caudell & Gallucci, 1995	Jemmott et al., 1983	Nagabhushan et al., 2001	Sabioncello et al., 2000		Kubitz et al., 1986	Song et al., 1999
Chi et al., 1993	Kamei et al., 1997, 1998	Pettingale et al., 1994	Scanlan et al., 1998		Leserman et al., 1997	Theorell et al., 1990
S. Cohen et al., 2000	Kang et al., 1996, 1997, 1998	Solomon et al., 1997	Schlesinger & Yodfat, 1988		Levy et al., 1989	Tjemsland et al., 1997
Cruse et al., 1993	Kiecolt-Glaser et al., 1986, 1993, 1994, 1997, 2001	Spratt & Denney, 1991	Stowell et al., 2001		Liang et al., 1997	Værnes et al., 1991
Delahanty et al., 1996, 1998, 2000	Kugler et al., 1996	Udelman, 1982	Vedhara et al., 1999		B. S. Linn et al., 1988	Vitaliano et al., 1998
Dopp et al., 2000	Lacey et al., 2000	Weiss et al., 1996	Vitaliano et al., 1998		M. W. Linn et al., 1983, 1984	Wilcox et al., 2000

Acute time-limited	Brief naturalistic	Event sequence	Chronic	Distant	Life event	Stress appraisal
Dugué et al., 1993 Endresen et al., 1991 Geenen et al., 1998 Gerits & DeBrabander, 1999 Gerritsen et al., 1996 Goebel & Mills, 2000 Goebel et al., 2000 Herbert et al., 1994 Jacobs et al., 2001 Jern et al., 1989 Johnson et al., 1996 Kamei et al., 1998 Kang & Fox, 2000 Landmann et al., 1984 Larson et al., 2001 Manuck et al., 1991 Marsland et al., 1995, 1997, 2001 Matthews et al., 1995 McDonald & Yagi, 1960 Miller, Dopp, et al., 1999 Mills & Dimsdale, 1996 Mills, Berry, et al., 1995 Mills et al., 1996, 1998 Mills, Haeri, & Dimsdale, 1995 Mills, Ziegler, et al., 1995 Moyna et al., 1999 Naliboff et al., 1991 Naliboff, Solomon, Gilmore, Benton, et al., 1995 Naliboff, Solomon, Gilmore, Fahey, et al., 1995 Neumann & Chi, 1999 Neumann et al., 1998, 2000 Ohira et al., 1999 Olff et al., 1995 Pawlak et al., 1999, 2000 Pehlivanođlu et al., 2001 Peters et al., 1999 Pike et al., 1997 Redwine et al., 2001 Ring et al., 2000 Rohleder et al., 2001 Sauer et al., 1995 Schedlowski, Jacobs, Alker, et al., 1993 Schedlowski, Jacobs, & Stratmann, et al., 1993 Schmid-Ott et al., 1998, 2001 Sgoutas-Emch et al., 1994 Sieber et al., 1992 Spangler, 1997 Stone et al., 1993 Tsopanakis & Tsopanakis, 1998 Uchino et al., 1995	Lowe et al., 2000 Maes et al., 1997, 1998, 1999 Marchesi et al., 1989 Marshall et al., 1998 Marucha et al., 1998 McClelland et al., 1985 Ockenfels et al., 1994 Paik et al., 2000 Segerstrom, 2001 Segerstrom et al., 1998 Song et al., 1999 Uchakin et al., 2001 Van Rood et al., 1995 Vassend & Halvorsen, 1987 Vedhara & Nott, 1996 Wadee et al., 2001 Whitehouse et al., 1996 Wolf et al., 1994 Workman & La Via, 1987	Zisook et al., 1994			Martin & Dobbin 1988 McClelland et al., 1980 McDade et al., 2000 McIntosh et al., 1993 McNaughton et al., 1990 Miletic et al., 1996 H. Moss et al., 1998 R. B. Moss et al., 1989 Mulder et al., 1995 Patterson et al., 1995 Perry et al., 1992 Petrey et al., 1991 Rabkin et al., 1991 Ravindran et al., 1996 Schlesinger & Yodfat, 1991 Shea et al., 1991 Thomason et al., 1996 Thornton et al., 2000 Vialettes et al., 1989 Zautra et al., 1989	

Acute time-limited	Brief naturalistic	Event sequence	Chronic	Distant	Life event	Stress appraisal
Van der Pompe et al., 1997, 1998						
Van der Voort et al., 2000						
Wang et al., 1998						
Weisse et al., 1990						
Willemsen et al., 1998						
Winzer et al., 1999						
Zakowski, 1995						
Zakowski et al., 1992, 1994						
Zeier et al., 1996						

Table 3

Meta-Analysis of Immune Responses to Acute Time-Limited Stress in Healthy Participants

Immune marker	*k*	*N*	*r*	SE_r	95% CI	*p*	*Q*
Leukocyte subset count							
Leukocytes	25	1,129	.17	.04	.10, .25	.001	34.61
Granulocytes	12	397	.08	.06	minus;.04, .19	.18	31.77
Neutrophils	3	86	.30	.12	.08, .50	.009	2.13
Eosinophils	3	81	−.10	.16	−.39, .21	.53	2.99
Monocytes	15	590	.04	.05	−.05, .13	.43	15.43
Lymphocytes	24	828	.18	.05	.09, .26	.001	31.77
T lymphocytes	33	1,452	.07	.03	.01, .12	.01	25.48
T-helper lymphocytes	42	1,678	.01	.03	−.05, .05	.86	23.72
T-cytotoxic lymphocytes	42	1,678	.20	.03	.15, .25	.001	34.05
T-helper:T-cytotoxic ratio	19	920	−.23	.10	−.40, −.04	.02	17.98
Naive T lymphocytes	3	241	−.09	.11	−.29, .12	.41	2.46
B lymphocytes	18	739	−.07	.04	−.14, .01	.08	16.23
Activated B lymphocytes	4	60	−.15	.14	−.40, .14	.31	0.48
Natural killer cells	41	1,635	.43	.06	.33, .51	.001	172.75***
Large granular lymphocytes	8	362	.53	.30	.00, .83	.05	165.64***
Leukocyte subset percentage							
Granulocytes	5	295	−.13	.10	−.31, .07	.20	7.24
Neutrophils	5	217	.04	.07	−.10, .18	.56	3.75
Monocytes	7	277	.06	.09	−.12, .23	.55	10.82
Lymphocytes	7	350	.06	.06	−.05, .16	.30	1.34
T lymphocytes	10	497	−.05	.09	−.22, .13	.62	28.05***
T-helper lymphocytes	14	642	−.24	.04	−.31, −.16	.001	13.61
T-cytotoxic lymphocytes	15	692	.09	.04	.01, .16	.03	9.28
B lymphocytes	5	248	−.11	.07	−.24, .02	.09	1.46
Natural killer cells	15	693	.24	.11	.03, .42	.02	90.19***
Total immunoglobulins							
Serum IgA	4	91	.12	.11	−.10, .33	.30	0.95
Serum IgM	3	67	.14	.13	−.12, .37	.30	0.61
Secretory IgA secretion rate	6	293	.22	.08	.06, .37	.008	6.92
Secretory IgA concentration	8	337	.22	.09	.05, .38	.01	13.05
Basal cytokine levels							
Interleukin-1β	4	89	−.01	.11	−.23, .21	.91	0.25
Natural killer cell function							
Natural killer cell cytotoxicity	37	1,398	.30	.05	.20, .39	.001	108.85***
Per-cell cytotoxicity	8	287	.12	.11	−.09, .32	.26	18.12*
Lymphocyte proliferation							
Proliferation to ConA	17	706	−.17	.04	−.24, −.09	.001	14.12
Proliferation to PHA	26	1,120	−.17	.04	−.23, −.10	.001	35.36
Proliferation to PWM	10	480	−.10	−.05	−.19, −.01	.03	5.84
Cytokine production							
Interleukin-1β	3	78	.01	.12	−.23, .23	.98	5.78
Interleukin-4	3	136	−.19	.11	−.39, .03	.08	2.38
Interleukin-6	3	143	.28	.09	.13, .44	.001	12.84**
Interferon-γ	3	96	.21	.11	.01, .40	.05	0.24

Note. CI = confidence interval; IgA = immunoglobulin A; IgM = immunoglobulin M; ConA = concanavalin A; PHA = phytohemagglutinin; PWM = pokeweed mitogen.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 4

Meta-Analysis of Immune Responses to Brief Naturalistic Stress in Healthy Participants

Immune marker	*k*	*N*	*r*	SE_r	95% CI	*p*	*Q*
Leukocyte subset count							
Leukocytes	9	249	.20	.07	.07, .32	.002	12.95
Granulocytes	3	56	.01	.15	−.27, .29	.93	0.01
Neutrophils	5	103	.11	.11	−.07, .34	.18	2.33
Monocytes	6	120	.06	.10	−.13, .25	.52	3.90
Lymphocytes	9	236	.06	.08	−.10, .23	.46	10.46
T lymphocytes	5	110	.03	.10	−.18, .22	.81	0.05
T-helper lymphocytes	7	197	.06	.08	−.09, .21	.43	1.08
T-cytotoxic lymphocytes	6	185	.05	.08	−.10, .20	.50	1.74
T-helper:T-cytotoxic ratio	12	351	.01	.07	−.11, .14	.84	13.68
B lymphocytes	5	126	.48	.56	−.51, .92	.35	99.48***
Natural killer cells	5	103	−.15	.11	−.35, .06	.16	2.06
Leukocyte subset percentage							
Monocytes	4	98	.11	.11	−.10, .32	.30	2.33
Lymphocytes	3	97	−.13	.11	−.33, .08	.23	2.05
T lymphocytes	5	160	−.16	.18	−.47, .19	.36	13.67**
T-helper lymphocytes	11	350	−.11	.10	−.29, .09	.28	26.56**
T-cytotoxic lymphocytes	12	362	−.03	.06	−.14, .08	.60	8.84
B lymphocytes	3	121	.07	.53	−.74, .80	.89	42.48***
Natural killer cells	5	163	−.02	.19	−.38, .35	.93	18.20**
Total immunoglobulins							
Serum IgA	6	243	.11	.07	−.02, .24	.10	1.28
Serum IgG	7	290	.06	.06	−.06, .17	.37	2.54
Serum IgM	7	290	.02	.10	−.17, .21	.83	13.41*
Secretory IgA rate	4	139	.09	.33	−.50, .63	.78	31.31***
Secretory IgA concentration	9	350	.19	.18	−.20, .46	.40	66.97***
Specific immunoglobulin							
Epstein-Barr virus	7	359	.20	.04	.10, .30	.001	6.56
Herpes simplex virus	4	225	.18	.08	−.02, .34	.08	4.97
Complement molecule							
C3	3	116	−.16	.10	−.34, .03	.09	1.77
Natural killer cell function							
Natural killer cell cytotoxicity	14	468	−.11	.05	−.21, −.01	.04	14.55
Lymphocyte proliferation							
Proliferation to ConA	9	220	−.32	.15	−.56, −.03	.03	27.08***
Proliferation to PHA	14	443	−.19	.09	−.35, −.02	.03	33.38***
Proliferation to PWM	3	106	−.17	.15	−.43, .12	.24	4.75
Cytokine production							
Interleukin-1β	6	149	.11	.08	−.05, .27	.17	15.07***
Interleukin-2	4	107	−.17	.36	−.71, .49	.63	27.34***
Interleukin-4	3	81	−.10	.12	−.32, .13	.39	0.69
Interleukin-6	3	100	.26	.11	.06, .44	.01	0.79
Interleukin-10	3	95	.41	.11	.21, .57	.001	1.65
Interferon-γ	8	314	−.30	.13	−.51, .05	.02	28.76***
Tumor necrosis factor-α	3	100	.18	.19	−.19, .51	.34	5.10

Note. CI = confidence interval; IgA = immunoglobulin A; IgG = immunoglobulin G; IgM = immunoglobulin M; ConA = concanavalin A; PHA = phytohemagglutinin; PWM = pokeweed mitogen.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 5

Meta-Analysis of Immune Responses to Brief Naturalistic Stress in Participants With Asthma

Immune marker	*k*	*N*	*r*	SE_r	95% CI	*p*	*Q*
Neutrophil function							
Superoxide release with FMLP	3	216	.20	.07	.06, .32	.004	0.39
Superoxide release with PHA	3	216	.37	.07	.24, .49	.001	0.68
Natural killer cell function							
Natural killer cell cytotoxicity	3	216	−.33	.07	−.45, −.21	.001	0.50
Lymphocyte proliferation							
Proliferation to PHA	3	216	.32	.07	.19, .43	.001	0.35

Note. CI = confidence interval; FMLP = *N*-formyl-met-leu-phe; PHA = phytohemagglutinin.

Table 6

Meta-Analysis of Immune Responses to Stressful Event Sequences in Healthy Participants

Immune marker	*k*	*N*	*r*	SE_r	95% CI	*p*	*Q*
Leukocyte subset count							
Monocytes	3	113	−.02	.10	−.21, .17	.87	0.39
Lymphocytes	5	223	.05	.07	−.09, .18	.49	2.65
T lymphocytes	5	213	−.02	.07	−.16, .12	.82	0.37
T-helper lymphocytes	9	566	.03	.11	−.19, .25	.81	39.29***
T-cytotoxic lymphocytes	8	544	−.14	.15	−.41, .15	.35	58.22***
T-helper:T-cytotoxic ratio	6	296	.06	.08	−.09, .21	.44	7.54
B lymphocytes	5	185	.02	.08	−.13, .17	.76	0.35
Natural killer cells	4	370	.17	.09	.00, .34	.05	5.06
Leukocyte subset percentage							
T lymphocytes	3	129	.02	.09	−.16, .19	.85	0.11
T-helper lymphocytes	5	279	.00	.06	−.12, .12	.94	0.00
T-cytotoxic lymphocytes	5	279	−.05	.06	−.17, .07	.43	3.65
B lymphocytes	3	129	−.04	.09	−.22, .14	.67	0.57
Specific immunoglobulin							
Epstein-Barr virus	3	198	.21	.07	.07, .34	.003	1.18
Natural killer cell function							
Natural killer cell cytotoxicity	13	698	−.03	.17	−.29, .34	.87	164.40***
Lymphocyte proliferation							
Proliferation to ConA	6	297	−.04	.06	−.15, .08	.53	2.53
Proliferation to PHA	11	675	.10	.10	−.09, .28	.32	42.25***
Proliferation to PWM	7	284	.12	.16	−.19, .40	.45	28.72***

Note. CI = confidence interval; ConA = concanavalin A; PHA = phytohemagglutinin; PWM = pokeweed mitogen.

*** $p < .001$.

Table 7

Meta-Analysis of Immune Responses to Chronic Stress in Healthy Participants

Immune marker	*k*	*N*	*r*	SE_r	95% CI	*p*	*Q*
Leukocyte subset count							
Leukocytes	4	240	.07	.07	−.06, .19	.32	2.12
Neutrophils	3	124	.36	.36	−.33, .79	.31	20.45***
Eosinophils	3	124	−.07	.22	−.47, .35	.75	8.07*
Monocytes	4	240	−.04	.17	−.36, .29	.83	14.33**
Lymphocytes	4	240	−.06	.10	−.25, .13	.54	5.24
T lymphocytes	5	470	−.03	.05	−.12, .06	.55	2.75
T-helper lymphocytes	10	786	−.05	.04	−.12, .03	.22	8.54
T-cytotoxic lymphocytes	10	786	−.08	.08	−.23, .08	.34	33.44***
T-helper:T-cytotoxic ratio	6	528	−.11	.08	−.29, .08	.26	17.47**
Activated B lymphocytes	3	138	−.02	.09	−.19, .15	.82	0.03
Natural killer cells	4	158	−.14	.32	−.65, .45	.65	33.61***
Leukocyte subset percentage							
Monocytes	3	224	.08	.10	−.11, .26	.42	3.18
T lymphocytes	5	522	−.03	.05	−.13, .07	.59	4.93
T-helper lymphocytes	10	860	−.07	.06	−.18, .03	.19	19.45*
T-cytotoxic lymphocytes	10	860	.02	.05	−.08, .11	.75	13.72*
Natural killer cells	6	246	.04	.09	−.13, .21	.64	7.85
Specific immunoglobulin							
Antibody to herpes simplex virus 1	3	185	.44	.34	−.19, .81	.17	20.78***
Antibody to influenza after vaccination	3	304	−.22	.05	−.33, −.11	.001	0.38
Natural killer cell function							
Natural killer cell cytotoxicity	8	563	−.12	.05	−.20, −.01	.04	11.58
Lymphocyte proliferation							
Proliferation to ConA	4	486	−.13	.06	−.24, −.02	.02	4.06
Proliferation to PHA	6	636	−.16	.06	−.27, −.05	.004	8.75
Cytokine production							
Interleukin-2	3	355	−.21	.05	−.31, −.11	.001	1.50

Note. CI = confidence interval; ConA = concanavalin A; PHA = phytohemagglutinin.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 8

Meta-Analysis of Immune Responses to Distant Stressors and Posttraumatic Stress Disorder in Healthy Participants

Immune marker	*k*	*N*	*r*	SE_r	95% CI	*p*	*Q*
Natural killer cell cytotoxicity	3	94	−.05	.25	−.49, .41	.84	7.67*

Note. CI = confidence interval.

* $p < .05$.

Table 9

Meta-Analysis of Immune Responses to Major and Minor Life Events of Unknown Duration in Healthy Participants

Immune marker	*k*	*N*	*r*	SE_r	95% CI	*p*	*Q*
Leukocyte subset count							
Lymphocytes	5	537	−.18	.17	−.47, .14	.27	20.28***
T lymphocytes	4	237	.00	.07	−.13, .13	.99	0.00
T-helper lymphocytes	5	227	.00	.07	−.13, .13	.99	0.00
T-cytotoxic lymphocytes	5	227	.05	.07	−.09, .18	.48	3.02
T-helper:T-cytotoxic ratio	3	70	.14	.38	−.54, .71	.71	12.11**
Natural killer cells	4	194	−.08	.07	−.22, .07	.28	2.72
Leukocyte subset percentage							
T lymphocytes	3	151	.20	.21	−.21, .55	.34	7.61*
T-helper lymphocytes	7	285	.01	.06	−.11, .13	.83	0.54
T-cytotoxic lymphocytes	6	205	−.01	.07	−.15, .14	.92	0.07
Natural killer cells	5	261	.00	.06	−.12, .12	.99	0.00
Total immunoglobulins							
Serum IgA	3	124	−.07	.10	−.26, .14	.52	2.19
Serum IgG	3	124	−.06	.10	−.24, .13	.54	2.06
Serum IgM	3	124	.03	.09	−.15, .21	.72	0.72
Secretory IgA rate	3	276	−.08	.10	−.26, .11	.43	3.97
Secretory IgA concentration	4	101	−.16	.14	−.42, .12	.25	4.34
Specific immunoglobulin							
Epstein-Barr virus	3	317	−.02	.11	−.23, .19	.86	5.65
Natural killer cell function							
Natural killer cell cytotoxicity	12	672	−.07	.07	−.20, .07	.35	29.39***
Lymphocyte proliferation							
Proliferation to ConA	3	72	−.13	.15	−.35, .16	.38	2.49
Proliferation to PHA	4	131	−.26	.15	−.50, .03	.08	6.11

Note. CI = confidence interval; IgA = immunoglobulin A; IgG = immunoglobulin G; IgM = immunoglob-ulin M; ConA = concanavalin A; PHA = phytohemagglutinin.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 10

Meta-Analysis of Immune Responses to Major and Minor Life Events of Unknown Duration in Participants With HIV/AIDS

Immune marker	*k*	*N*	*r*	SE_r	95% CI	*p*	*Q*
Leukocyte subset count							
T-helper lymphocytes	11	998	−.01	.03	−.08, .05	.70	7.70
T-cytotoxic lymphocytes	6	669	−.14	.08	−.29, .01	.08	17.92**
T-helper:T-cytotoxic ratio	3	356	−.02	.05	−.13, .09	.70	0.09
Natural killer cells	3	261	−.27	.06	−.38, −.15	.001	0.30
Leukocyte subset percentage							
T-helper lymphocytes	4	1,026	−.02	.06	−.15, .10	.73	7.58
T-cytotoxic lymphocytes	3	223	.00	.07	−.13, .13	.99	0.00

Note. CI = confidence interval.

** $p < .01$.

Table 11

Meta-Analysis of Immune Responses to Global Stress Appraisals in Healthy Participants

Immune marker	*k*	*N*	*r*	SE_r	95% CI	*p*	*Q*
Leukocyte subset count							
T lymphocytes	3	241	−.15	.09	−.31, .03	.10	3.15
T-helper lymphocytes	3	241	−.14	.10	−.32, .06	.18	3.80
T-cytotoxic lymphocytes	4	279	−.02	.09	−.19, .15	.80	5.09
Naive T lymphocytes	3	241	−.09	.11	−.29, .12	.41	4.29
Natural killer cells	3	205	−.20	.13	−.42, .04	.10	4.28
Leukocyte subset percentage							
T-helper lymphocytes	3	143	−.02	.09	−.19, .15	.79	0.08
T-cytotoxic lymphocytes	3	143	−.03	.09	−.23, .11	.48	0.60
Total immunoglobulin							
Serum IgG	4	332	.02	.10	−.18, .20	.87	7.51
Natural killer cell function							
Natural killer cell cytotoxicity	4	151	−.11	.09	−.27, .06	.21	1.85

Note. CI = confidence interval; IgG = immunoglobulin G.

Table 12

Meta-Analysis of Immune Responses to Stress Appraisals and Intrusive Thoughts Within Healthy Stressed Populations

Immune marker	*k*	*N*	*r*	SE_r	95% CI	*p*	*Q*
Leukocyte subset count							
T-helper lymphocytes	3	462	−.10	.11	−.31, .11	.35	7.52*
T-cytotoxic lymphocytes	3	462	−.26	.32	−.71, .34	.40	57.99***
Natural killer cell function							
Natural killer cell cytotoxicity	3	566	−.15	.06	−.27, −.02	.02	7.97

Note. CI = confidence interval.

* $p < .05$.

*** $p < .001$.

frontiers in Psychology

REVIEW
published: 14 December 2018
doi: 10.3389/fpsyg.2018.02584

The Developmental Origins of the Social Brain: Empathy, Morality, and Justice

Chenyi Chen[1,2,3,4†], *Róger Marcelo Martínez*[1,5†] *and Yawei Cheng*[1,5,6*]

[1] Department of Physical Medicine and Rehabilitation, National Yang-Ming University Hospital, Yilan, Taiwan, [2] Graduate Institute of Injury Prevention and Control, College of Public Health, Taipei Medical University, Taipei, Taiwan, [3] Research Center of Brain and Consciousness, Shuang Ho Hospital, Taipei Medical University, New Taipei City, Taiwan, [4] Institute of Humanities in Medicine, Taipei Medical University, Taipei, Taiwan, [5] Institute of Neuroscience and Brain Research Center, National Yang-Ming University, Taipei, Taiwan, [6] Department of Education and Research, Taipei City Hospital, Taipei, Taiwan

The social brain is the cornerstone that effectively negotiates and navigates complex social environments and relationships. When mature, these social abilities facilitate the interaction and cooperation with others. Empathy, morality, and justice, among others, are all closely intertwined, yet the relationships between them are quite complex. They are fundamental components of our human nature, and shape the landscape of our social lives. The various facets of empathy, including affective arousal/emotional sharing, empathic concern, and perspective taking, have unique contributions as subcomponents of morality. This review helps understand how basic forms of empathy, morality, and justice are substantialized in early ontogeny. It provides valuable information as to gain new insights into the underlying neurobiological precursors of the social brain, enabling future translation toward therapeutic and medical interventions.

Keywords: empathy, morality, justice, interpersonal harm aversion, inequity aversion

OPEN ACCESS

Edited by:
Ilaria Grazzani, Università degli Studi di Milano-Bicocca, Italy

Reviewed by:
Elisabetta Conte, Università degli Studi di Milano-Bicocca, Italy
Peter J. Reschke, Brigham Young University, United States

***Correspondence:**
Yawei Cheng
ywcheng2@ym.edu.tw

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *18 September 2018*
Accepted: *03 December 2018*
Published: *14 December 2018*

Citation:
Chen C, Martínez RM and Cheng Y (2018) The Developmental Origins of the Social Brain: Empathy, Morality, and Justice. Front. Psychol. 9:2584. doi: 10.3389/fpsyg.2018.02584

INTRODUCTION

The ability to detect contextual cues, such as those of distress and need, has been proposed as the genesis of the mechanism by which basic forms of mimicry and conditioning will naturally evolve into empathy, and later result in helping/moral behaviors (Hoffman, 2007). The perceptual capacity for identifying such cues is significantly meaningful in evolutionary terms, both in regards to the parental care of the offspring, as well as for intra-kin group bonding purposes, as it facilitates group survival. The basic affective circuits on which this ability rests upon emerge much earlier in the brain's evolution than do higher cognitive capacities (Decety and Svetlova, 2012). The underlying mechanisms that enable mammalian species to discern and react with care to the distress and suffering of another, derive from evolutionarily archaic subcortical circuits (e.g., brainstem, amygdala, hypothalamus, and basal ganglia) and neuro-hormonal systems and processes interrelated with attachment, parental care, and affective sensitivity (Tucker et al., 2005; Decety et al., 2016).

Moral behavior has been theorized to have its roots in this distress detection ability. When human infants perceive distressful cues derived from the other, they themselves will also be affected by this negative emotionality as if it was their own, this due to the children's premature brain

still being unable to effectively differentiate between self and other. Consequently, the infants will try suppressing this distress, but only as a mean to overcome the distress that is produced and mirrored within themselves. This primitive form of egocentric empathy is the precursor for moral behavior. But it is this archaic form of empathy the one that will, across time, and thanks to a future successful differentiation between self and other, turn into genuine, real empathy for the other's distress (Hoffman, 2007).

Actions to alleviate another infant's afflictions (e.g., by sharing toys with the other, or comforting him) in emotionally painful situations, are already forms of altruism. As such, these behaviors are pro-social and moral in nature. Altruism at the same time encompasses instrumental helping, whereby an infant acts charitably on behalf of another to help him achieve his goal. But behaviors such as helping and sharing resources inevitably lead to notions such as justice (Warneken and Tomasello, 2009a,b).

The aim of this review is to give account of the neural processes underlying each of the aforementioned aspects of the social brain, beginning with the distress detection ability, going through how this relates to empathy and harm sensitivity, thus, impacting moral behaviors (e.g., inequity aversion), and ending with how the typical processing of others' distress is affected in different neuropsychiatric disorders, consequently conditioning pro-social behaviors. By the end of this paper, our intention is to have effectively bridged some of the existing psychological and neuroscientific research dealing with the themes of empathy, morality, and justice.

PERCEIVING SIGNALS OF DISTRESS AND NEED ACROSS SPECIES

Bird and mammal progenitors who learn to be affected by their offspring's needs, as a function of their ability to detect such signals, are able to secure the survival of their offspring more successfully than those who remain indifferent due to a lack of such capacity. Consequently, and unsurprisingly enough, a communication system would develop over time and across species where the children's stylized distress signals would automatically trigger parental care, an ability which later would evolve into even more complex mechanisms such as perspective taking and empathy (de Waal, 2008). Likewise, emotion contagion is an archaic and basic form of hereditary intra-species communication previously observed and researched in many vertebrates (Hatfield et al., 1993). In one study, rats that were taught to action a lever in order to receive food would refrain from doing so if their actions were coupled with the delivery of electrical shocks to an unconcealed, adjacent rat (Church, 1959). Therefore, we can infer that rats perceive a conspecific's pain as aversive. In another study, rats that were allowed to run free learned to intentionally and quickly open the restrainers of their cage-mates, in order to also liberate them. In another experiment of this same study, when releasing a cage-mate was pitted against obtaining chocolate placed inside a second cage, rats would typically open both of the restrainers, and share the acquired chocolate with said cage-mate (Ben-Ami Bartal et al., 2011). The finding that rats show pro-social behaviors as a result of a conspecific's distress provides compelling proof for the notion that empathy and helping behavior have their roots in biology. Moreover, the rats' helping behavior was observed to be reduced by anxiolytic treatment, and associated with sympatholytic corticosterone response (Ben-Ami Bartal et al., 2016). These findings suggest that affect has a fundamental, motivational role in the pro-social behavior of rodents, as shown by the necessity of the rat-helper to resonate with the affective status of the rat-victim.

Nevertheless, there is a downside. Because empathy evolved within the confines of parental-care and cooperatively group coexisting, there's a great probability that it may still be biased and limited to benefit only "known" or "in-group" members, as members of other groups can be seen as competitors or free-riders (Tomasello and Gonzalez-Cabrera, 2017). For example, it was observed that rats did not help other rats of a different strain, as they were considered strangers. This would change, however, if the freed rat was previously housed with the trapped one. Thus, pair-housing rats would prompt them to help each other, even when they belonged to a different strain (Ben-Ami Bartal et al., 2014). This outcome may provide evidence to support the notion that rats can broaden their pro-social motivational scope to include phenotypically similar others. Consistent with kin selection and reciprocal altruism theory, individuals regarded as familiar, as well as previous cooperators, are favored from an empathy-related standpoint (de Waal, 2008). As stated before, the drawback is the possibility that empathy may not lead all the time and invariably toward morality -as these findings let us observe, and can sometimes be the starting point for wrongful actions by championing exclusively for kin- or self-related interests (Decety and Cowell, 2014b).

DISTRESS PERCEPTION IN THE HUMAN INFANT BRAIN

At the very early stages of ontogeny in humans, the neonatal brain has been found to have enough sensitivity to discriminate distressful and threat-related voices from emotionally neutral sounds, even when the neonates were in a sleeping state (Cheng et al., 2012b; Zhang et al., 2014).

Lloyd-Fox et al. (2012) demonstrated that 4- to 7-month-old infants already possess the ability to differentiate between voices (e.g., coughing, yawning, throat clearing, laughing, and crying) and non-vocal sounds (e.g., sounds from water running and toys). Moreover, Cheng et al. (2012b) observed that when newborns discern between the emotional voices and the non-vocal sounds, their responses are dominant at the right hemisphere, which is the brain region related to natural speech perception (Alexandrou et al., 2017). In one functional magnetic resonance imaging (fMRI) study, neutral voices (e.g., coughing, sneezing, yawning, and throat clearing), happy voices (laughing), sad voices (crying), and non-vocal sounds (toy and water running sounds) were presented to 3- to 7-month-old babies (Blasi et al., 2011). Infants at this age were found to have activations not only in voice-sensitive regions (the right anterior middle and superior temporal gyri, and medial frontal gyri), but also in brain regions specific

to negative voices (i.e., the left orbitofrontal cortex and insula). Since it has been observed that the ability of speech perception, especially for discriminating phonetic contrasts, emerges during the first month of birth (Dehaene-Lambertz and Dehaene, 1994; Cheour et al., 1998; Dehaene-Lambertz and Pena, 2001; Kuhl, 2004), and because of the fact that voice perception emerges earlier than speech perception in human development (Belin et al., 2004; Blasi et al., 2011), it is plausible to conclude that the ability to perceive voices appears before, or at least around birth. This conclusion is supported by previous studies yielding findings that demonstrate that fetal heart rate increases when hearing the maternal voice, and decreases when listening to a stranger's (Kisilevsky et al., 2003). These results suggest that fetuses already possess vocal identity. In the same manner, Beauchemin et al. (2011) reported that newborns could discern the mother's voice from a female stranger's one. Furthermore, according to Cheng et al. (2012b) serial experiments, newborns were sensitive to emotional voices beyond specific language, but it is key to note that the specificity is driven by voice perception *per se*. Thus, it sounds reasonable to infer that newborns exhibit the ability comparable with adults to process the affective information, such as pain and distress, being conveyed through the voice (Schirmer et al., 2005; Fan et al., 2013; Chen et al., 2014, 2015, 2016b, 2017a; Hung and Cheng, 2014). Voices with affective information are presumed to elicit more processing resources than those without affect. Hence, the emergence of the specialization for processing emotions is already advanced at the first days of life.

When visual affective discrimination is involved, studies have shown that between 18 and 24 months of age, and comparable to auditory stimuli, infants show decreased ERPs when exposed to strangers' faces than when exposed to their mother's face (Carver et al., 2003). Nevertheless, children's initial ability to discriminate facial emotions is not as effective, and only significantly improves with age (Boyatzis et al., 1993). Children's attention to visual cues varies depending on if the stimuli is static or dynamic. The infants' processing of visual whole-body static images (whether their faced is veiled or not) appears to be immature, but when shown dynamic whole-body stimuli (a more real-life-like scenario), their processing resembles that of adults (Nelson and Mondloch, 2018).

When it comes to the neural underpinnings of visual affective cues processing, a recent review paper by Bachmann et al. (2018) reports that emotional stimuli evokes greater neural responses than neutral stimuli, and that the brain regions observed to have a significant activation during these processes encompass the action observation network (conformed by the inferior frontal gyrus, the premotor cortex, and the inferior parietal lobe), suggesting that human brains interpret others' actions by employing motor simulations grounded on the motor programs they already possess (Kilner et al., 2007); the mentalizing network (which includes the temporo-parietal junction, the temporal pole, the lateral orbitofrontal cortex, and the dorsomedial prefrontal cortex), making it possible for the individual to infer the other's mental state (also known as theory of mind) (Frith and Frith, 2006); as well as other regions whose task is to process body motion and form, such as the extra striate body area, the fusiform body area, and the posterior superior temporal sulcus, which interestingly respond selectively to emotional body movements (Downing et al., 2001; Blake and Shiffrar, 2007; Peelen et al., 2007; Pichon et al., 2009; Sinke et al., 2010; Kret et al., 2011; Atkinson et al., 2012). In addition, the amygdala and the hypothalamus have also been observed as having pivotal roles when observing body expressions containing emotional valences, as the former brain region is involved in emotional processing, while the latter holds a crucial role in regards to defensive reactions and in action preparation (Barbas et al., 2003). What's more, due to self-protection purposes, an appropriate recognition and a suitable response to threatening or dangerous individuals is of utter importance, thus, the processing of this experiences is particularly favored. This is suggested by the fact that emotional content possessing the aversive valences of threat and anger, in contrast to those containing negative valences of fear and sadness, specifically elicit greater activity in brain structures such as the superior temporal sulcus, the premotor cortex, the temporo-parietal junction, the dorsomedial prefrontal cortex, and the lateral orbitofrontal cortex (Pichon et al., 2008, 2009; Peelen et al., 2010).

Bachmann et al. (2018) further posit that when it comes to attentional matters, implicit and explicit tasks appear to activate brain regions differentially, while the hypothalamus, the premotor cortex, and the amygdala are equally activated by both implicit and explicit tasks, other areas, in particular those conforming the mentalizing network, react to explicit judgements of emotionally charged stimuli.

AFFECTIVE AROUSAL/EMOTIONAL SHARING

The effective perception of distress will give rise to several other facets involved in human empathy, including affective arousal/emotional sharing, empathic concern, and perspective taking. Each facet of empathy affects moral cognition differently and predicts distinct consequences regarding moral behavior (Decety and Cowell, 2014a). Among them, the affective element comprising empathy evolves earlier than the cognitive aspect. For instance, newborns and infants become significantly distressed immediately after another newborn starts crying (Dondi et al., 1999). What's more, they themselves also begin crying when this happens, and their cry shows as much distress as that of the infant who initiated it. Noteworthy, is that this cry is far from just a form of imitation, as it is not as intense when they are exposed to a chimpanzee's cry, nor even when they are exposed to a recording of their own cry (Simner, 1971; Sagi and Hoffman, 1976; Martin and Clark, 1982). Furthermore, human infants (around 10 weeks of age) can readily imitate expressions of fear, sadness, and surprise (Haviland and Lelwica, 1987), fitting individuals for future empathy-based connections by means of affective interactions with others (Decety, 2010). Notwithstanding, when it comes to this form of imitation, a recent longitudinal study put into question the notion that this ability is present since birth (Oostenbroek et al., 2016). It was found that there was no such imitation by neonates, attributing previous findings to be the cross-sectional nature of the research.

When it comes to emotional sharing, this capacity may be partially underpinned by the mirror neuron system (MNS), which -as observed by electroencephalographic (EEG) studies- appears to be already operating in infants with ages as young as 6 months old (Nystrom, 2008). First observed in the 1990's in non-human primates, the MNS is a system hypothesized to be the motor behind action understanding and imitation (Lepage and Theoret, 2007), hence, having an important role in perspective-taking and empathy (Nystrom, 2008), assumptions which have drawn much debate in the scientific community. While some research have found evidence just partially supporting these hypotheses (Dinstein et al., 2008), others have argued against these claims to the extent that they even doubt the existence of an MNS in humans (Hickok, 2009), although more recent studies have endorsed the initial assumptions (Woodward and Gerson, 2014), as well as supported the notion that the MNS in the human brain is already present since infancy. Furthermore, Meltzoff (2007) considered the MNS as a possible platform for the foundation of social cognition as a whole. He argued that infants effectively identify and match between actions performed by others and the proprioception arising from their own bodily movements. By registering these equivalences between the acts performed by another and those performed by the self, the infant is able to perceive that the others are 'like me.' This recognition of others as having similar perceptions and emotions would form the bedrock and starting point for social cognition. Nevertheless, further research is warranted as to elucidate the issues and intricacies, as well as to bridge the diverging literature, surrounding the MNS.

Moreover, in research using EEG/ERP in both children and adults, there is ample evidence showing that passive observation of visual stimuli depicting physical injuries to body parts, triggers an early component (N2, ∼200-ms) and a late-positive potential (LPP, ∼800-ms) (Chen et al., 2012; Cheng et al., 2014; Fan et al., 2014). The early N2 within a time window of 200 to 300-ms was observed to be modulated by attention to emotionally salient stimuli, and reflects affective arousal or emotional sharing; whereas LPP within 500 to 800-ms indexes cognitive reappraisal and emotion regulation (Li and Han, 2010). Adolescents, in comparison to young adults, elicited an earlier N2 in reaction to another's pain, and a greater LPP in response to neutral stimuli, demonstrating that affective and cognitive empathy is still developing during the adolescent years (Mella et al., 2012).

Throughout early development, children display enhanced N2 to pain during empathic concern engagement. Larger early N2 ERPs were evoked when perceiving painful stimuli versus when perceiving neutral stimuli. Within the slow wave window of LPP, greater differences between painful and neutral images occurred in the empathic concern condition when compared to the perspective-taking condition. Increased pain-evoked N2 responses in the cognitive empathy condition were also associated with empathy regarding parental disposition (Decety et al., 2018). Multilevel analyses, including neurophysiological responses to different empathic conditions and parental empathy levels, is warranted to demonstrate the important differences between the various aspects in children's empathy. This provides a thought-provoking link between parental dispositions and the children's neural workings during early development.

EARLY INTERPERSONAL HARM SENSITIVITY

One recent framework proposed the importance of third-party harm aversion as a necessary platform for constructing morality (Decety and Cowell, 2018). When presented with wooden characters pushing a red circle up (i.e., helping) or down (i.e., hindering) a hill, preverbal infants (6- and 10-month-olds) showed a negativity bias by staring significantly longer at, and reaching for the pro-social character, indicating aversion to anti-social behavior (Hamlin et al., 2010; Hamlin, 2014). In another experiment using the same paradigm, when the red circle was attached with "googly eyes" (social condition), 3-month-old infants preferred the character helping the climber uphill, over the character pushing the climber downhill, but they didn't show any preference when the "googly eyes" were not attached to the circle (inanimate control condition). The 3-month-olds' preference for the helpers over the hinderers appeared to indicate specific social evaluation (e.g., a perceptual predilection for uphill-helpers over downhill-pushers) (Hamlin et al., 2010). In addition, when researching on sympathy-related behaviors in 10-month-old infants, one study observed infants' predilection for victims, as opposed to a preference toward aggressors and neutral objects, after watching as a third-party social interactions containing enacted aggression. This indicates that preverbal infants exhibit sympathy-based responses toward others when they are attacked, even when they don't display any distress (Simner, 1971; Sagi and Hoffman, 1976). Furthermore, research has showed that 3-year-olds tend to avoid helping adults whose intentions are forthright harmful (Vaish et al., 2010), or tend to protest when observing this actions as a third-party (Vaish et al., 2011). Thus, a primitive form of sympathy toward others, emanating from interpersonal harm sensitivity that goes beyond being a simple response to distress as a consequence of emotional contagion, arises earlier during development than previously concluded (Kanakogi et al., 2013; Lee et al., 2015). In line with these findings, Scott and Baillargeon (2017) have pointed out that previous literature has misreported false-belief understanding, which empathy seemingly predicts (Ferguson et al., 2015), as emerging until the 4 years of age, this due to the experimental design taking into consideration traditional tasks, and excluding nontraditional ones (e.g., spontaneous-response and elicited-intervention tasks). When we take into account the latter, infants under 2 years of age have been seen to already exhibit false-belief understanding. Hence, and given the infants' premature verbal ability and insufficient information-processing resources (Scott and Roby, 2015), the aforementioned third-party moral evaluations can be considered to be driven by intuitive processes that form the understructure of a natural, untaught and unlearned moral core, molded by natural selection in order to facilitate social cohesion and cooperation (Hamlin, 2015).

Recently, research in developmental neuroscience that has taken into consideration parental dispositions has largely

confirmed these behavioral findings, and identifies specific neural mechanisms underpinning early socio-moral evaluations and their correspondence to moral preferences. In one study, infants (ages 12–24 months) watched dynamic visual stimuli depicting cartoon characters intentionally executing either pro-social or antisocial acts, while EEG, time-locked ERP, and gaze fixation were recorded. After watching the animation, the experimenters provided a physical version of the helper character and another of the hinderer, as to assess the infants' reaching preference. This yielded several specific neural outcomes, which include asymmetrical frontal spectral power densities, eye-tracking differences, and time-locked condition differences in the relatively automatic ERPs component (Nc, within 300 to 500-ms time window), at the time social evaluations were occurring. Children's reaching preference for the pro-social character, as well as parental sensitivity to injustice for others, could be predicted by the early automatic ERPs component, which were sensitive to the perception of pro-social versus antisocial scenarios (Cowell and Decety, 2015a). This differentiation should be understandably basic in essence, being firmly established in primitive resource allocation to relevant stimuli, and approach/withdrawal dispositions. The domain-general systems concerning attention and self-regulation, control early evaluations of social and moral nature, supporting increasing evidence yielded by the neuroscience of morality.

Another neurodevelopmental study that probed into implicit moral evaluations of antisocial and pro-social behaviors in children aged 3–5 years old showed distinct ERP components for early automatic attention (EPN) while they were watching characters enacting helping behaviors, and later cognitive control (LPP) while watching characters enacting harming behaviors. Noteworthy, later (LPP) waveforms, but not early (EPN) ones, predicted the children's current generosity (Cowell and Decety, 2015b). These findings further underscore that automatic, intuitive/affective, and cognitively controlled processes are required for what appear to be a basic third-party discernment concerning harming and helping behaviors.

INEQUITY AVERSION

Taking into account helping behaviors, a considerable amount of research suggests that the predecessors of what motivates justice have evolved and developed due to environmental pressures where stabilized group cooperation was needed for survival purposes (Decety and Yoder, 2017). Observational studies about non-human primates' natural interactions served as the best opportunity to look into these primates' understanding of fairness and justice. Some primate species living in unique social systems characterized by exceptional social permissiveness and cooperation (which go all the way from sharing food to child care) have evolved inequity aversion, consequently responding negatively when receiving less reward than their social partners (Brosnan, 2013). As such, these species exhibit enhanced pro-social behavior when compared to other primates. Thus far, the extant findings best back the hypothesis that an aversion to inequity arose jointly with cooperative behavior between unrelated individuals (Price and Brosnan, 2012).

In the studies with a violation-of-expectancy paradigm, devised to assess expectations in human infants regarding equal distribution of resources, 15-month-old infants showed increased attention to unfair (unequal) over fair (equal) outcomes (Sommerville et al., 2013). In another developmental study, 19-month-old children expected a distributor to divide resources equally between two similar individuals. This expectation could be attributed to the sensitivity toward social inequity, rather than to low-level sensory processing, given that it was absent when the individuals were replaced with inanimate objects. Furthermore, 21- to 36-month-olds have been observed to already expect equity in distribution when individual efforts resulting in cooperation are taken into consideration. Hence, infants expected that a hard worker and a slacker should not be rewarded equally (Sloane et al., 2012). Nevertheless, and interestingly enough, it has also been observed that during this age window, rewards are shared more equitably when there is cooperation than when there is just parallel-, non-cooperative work involved (Hamann et al., 2011). The sensitivity degree to fairness as a third-party observer in 15-month-old infants, was seen to be associated to whether they shared toys altruistically or selfishly, denoting that moral appraisal and pro-social behavior are highly interconnected since early in development (Schmidt and Sommerville, 2011). These other-regarding inclinations emerge in a parallel and close knitted manner. The origin of a basic sense of fairness and altruism during infancy plays a pivotal role in the development of human-specific types of cooperation.

However, even though research has seemingly led us in a straight line from empathy to distress relief behaviors, toward inequity aversion, and leading to fairness and justice, it is important to underscore that this bridge needs to be treaded carefully. Empathy can inevitably lead to moral, pro-social behaviors such as altruism (de Waal, 2008), and, as stated before, altruism is greatly associated with morality and fairness (Schmidt and Sommerville, 2011). Nevertheless, not all forms of fairness, such as that stemming from collaboration in which both actors reap rewards, is considered as pro-social or altruistic, due to its underlying selfish nature (Warneken and Tomasello, 2009a). Yet, we can still push past this issue. Research was conducted with 21- and 27-month-old infants, where they had to collaborate with an adult toward the achievement of a joint goal. At one moment during the experiment, the adult simulated that he was unwilling or unable to continue with the task. When the latter occurred, the children tried to re-engage their collaborator with the same magnitude on both scenarios: when they needed their help to reach their individual goal, and when they didn't need of their participation to do so, irrespectively, and in contrast to when the adult was unwilling to complete such task. Therefore, this study demonstrated that infants in the employed age range were not only capable of understanding the other person's intentions, but were also able to act altruistically toward them (what is referred to as instrumental helping), rather than just using the other as a mere social tool (Warneken et al., 2012). Due to these later findings, we can now go full circle from distress perception and empathy, passing by affective arousal and harm sensitivity, to

inequity aversion and fairness, and back again to empathy and empathically led pro-social behavior.

ATYPICAL EARLY/AUTOMATIC PROCESSING IN NEUROPSYCHIATRIC DISORDERS

One source of evidence concerning the crucial role of these primitive forms of the social brain, including automatic distressful voice perception, early affective arousal/emotional sharing, and early sensitivity to interpersonal harm during the development, comes from atypical socioemotional processing as a result of neuropsychiatric disorders.

Previous studies using a passive auditory oddball paradigm to investigate the automatic processes in human voice perception, revealed that people with autism spectrum disorder (ASD) manifest general impairments in affective voice discrimination, as well as in low-level acoustic distinction. In addition to decreased amplitudes of mismatch negativity (MMN), which is a reliable index of the neural representations underlying automatic central auditory perception in response to acoustically matched non-vocal sounds, people with ASD failed to discriminate between happy and angry syllables (thus, having an impaired distress detection ability). Weak amplitudes of angry-evoked MMN were associated to severe autistic traits. As a result, examining emotional MMN may offer an opportunity to facilitate an early diagnosis for infants at risk for ASD (Fan and Cheng, 2014). Early-onset conduct disorder (CD), the major childhood predecessor to antisocial personality disorder in adulthood, also showed an atypical fashion in the processing of distressful voices at the pre-attentive level. The presence of increased differences in MMN between fearful and sad voices, and correlations between MMN and impulsive tendencies in youths demonstrating CD symptoms and a history of delinquent behaviors, helps shed light on the neural mechanisms of aggression (Hung et al., 2013). Atypical neurophysiological responses to threatening voices were also evidenced in patients with chronic schizophrenia, suggesting general impairments of voice perception and acoustic discrimination. The emotional salience processing of voices exhibited atypical characteristics at the pre-attentive level, which were related with positive symptoms in schizophrenia. These results may present evidence for bottom-up (i.e., perceptually based) cognitive remediation strategies, and indicates that emotional MMN may be a potential neurophysiological endophenotype for schizophrenia (Chen et al., 2016a).

In one neuroimaging study, when visual scenarios depicting intentional and unintentional harm were presented to CD with callous unemotional traits, a reduced hemodynamic response was found in the insula, a brain region which plays a crucial role in empathy and emotional awareness. Additionally, a weaker activation of the right posterior insula, as induced by perceiving the harm of others, was associated with more CD symptoms and callousness. Similarly, the reactive aggression scores of CD were associated with hyperactivity in the posterior insula and anterior mid-cingulate cortex, indicating heightened emotional response to provocative stimuli (Michalska et al., 2016). Another study applying electroencephalography and event-related potentials (EEG/ERPs) in juvenile psychopaths, appears to echo with the aforementioned findings. Youths with higher callous unemotional traits exhibited atypical neural dynamics in regards to pain empathy processing in the early stages concerning affective arousal, which was coupled with their comparable insensitivity to actual pain. Their ability to understand others' intentions, however, was not seen to be affected. Such disassociation between affective arousal and emotion understanding might contribute to generating aggressive behaviors in juvenile psychopaths (Cheng et al., 2012a).

Importantly, the link between pain sensitivity and empathic response was uncovered in adolescents with autistic and CD symptoms (Chen et al., 2017b). When compared to typically developing controls, the tactile pressure pain threshold was lower in the autistic group when compared to the conduct group, whose threshold was higher. The autistic group exhibited decreased ratings of unpleasantness and pain intensity to the sufferings of others than did the controls and the conduct groups. In the autistic and conduct groups, pain intensity scores were significantly associated with unpleasantness ratings to others' pain. Moreover, the color autistic group significantly differed from the controls in the association between pain threshold values and unpleasantness scores. Importantly, these results may cast some light on the relationship between atypical low-level sensory functioning, for example altered pain sensitivity, and high-level processing of empathy.

CONCLUSION

This review sheds some light on the primitive forms of the social brain as observed in early ontogeny. To distinguish among the different facets of empathy, including affective arousal/emotional sharing, interpersonal harm sensitivity, and perspective taking, it is critical to avoid illustrating the complex relationship between morality and empathy in a misleading and equivocal way. The presence of affective arousal/emotional sharing, a basic form of empathy, helps in the understanding of such relations more precisely, as well as assists in elucidating the concepts being used in the current literature and, perhaps, also aids in the abandonment of the muddy concept of empathy (Decety and Cowell, 2014a). The interpersonal harm aversion exhibited in early life, shapes the fundamental nature of moral and social cognition, indicating that such motivational value lies in more general processes, rather than in fully distinct, specific neural regions dedicated for morality (Decety and Cowell, 2018). The realization of how basic forms of empathy, morality, and justice are substantialized in early ontogeny may provide valid information as to gain new insights into the underlying neurobiological precursors of the social brain, which may enable future translation toward therapeutic and medical interventions.

Likewise, suggestions for future lines of research can be inferred from this review. Foremost, we echo the concerns

voiced by other authors (Nelson and Mondloch, 2018) that new studies in the fields of distress- and emotion-perception in others should adopt experimental designs employing dynamic stimuli over static ones, as they more clearly resemble daily interpersonal interactions. Furthermore, research is warranted as to elucidate to what extent empathy in the human social brain is innately reserved toward kin- and self-related interests, as well as how strong is the role played by parental dispositions in the infants' emergence and maturation of empathic concern. Parallel to the last point, scholars have argued that collaboration and instrumental helping might be grounded on distinct psychological processes (Warneken and Tomasello, 2009a), thus, it would be of great avail to address this matter in conjunction with a neuroscientific standpoint, as to identify the neural underpinnings driving such differences. Additionally, new lines of enquiry should be initiated in order to tackle the issues (maybe even the existence) surrounding the MNS and its relationship to empathy and social cognition. Last but not least, we make a call for further studies resulting in the development of therapeutic interventions for those psychiatric and neurological conditions where perceptual processes underlying empathy and moral cognition have been seen to be affected in an atypical manner.

AUTHOR CONTRIBUTIONS

CC and YC conceived and designed the review. CC and RM did the literature collection. CC and YC wrote the first draft of the manuscript. All authors contributed to manuscript write-up.

FUNDING

The study was funded by the Ministry of Science and Technology (MOST 106-2420-H-010-004-MY2; 106-2410-H-010-002-MY2; 107-2314-B-038-012), National Yang-Ming University Hospital (RD2017-005), Taipei Medical University (TMU106-AE1-B32 and DP2-107-21121-01-N-03), and the Brain Research Center, National Yang-Ming University from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

REFERENCES

Alexandrou, A. M., Saarinen, T., Makela, S., Kujala, J., and Salmelin, R. (2017). The right hemisphere is highlighted in connected natural speech production and perception. *Neuroimage* 152, 628–638. doi: 10.1016/j.neuroimage.2017.03.006

Atkinson, A. P., Vuong, Q. C., and Smithson, H. E. (2012). Modulation of the face- and body-selective visual regions by the motion and emotion of point-light face and body stimuli. *Neuroimage* 59, 1700–1712. doi: 10.1016/j.neuroimage.2011.08.073

Bachmann, J., Munzert, J., and Kruger, B. (2018). Neural underpinnings of the perception of emotional states derived from biological human motion: a review of neuroimaging research. *Front. Psychol.* 9:1763. doi: 10.3389/fpsyg.2018.01763

Barbas, H., Saha, S., Rempel-Clower, N., and Ghashghaei, T. (2003). Serial pathways from primate prefrontal cortex to autonomic areas may influence emotional expression. *BMC Neurosci.* 4:25. doi: 10.1186/1471-2202-4-25

Beauchemin, M., Gonzalez-Frankenberger, B., Tremblay, J., Vannasing, P., Martinez-Montes, E., Belin, P., et al. (2011). Mother and stranger: an electrophysiological study of voice processing in newborns. *Cereb. Cortex* 21, 1705–1711. doi: 10.1093/cercor/bhq242

Belin, P., Fecteau, S., and Bedard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends Cogn. Sci.* 8, 129–135. doi: 10.1016/j.tics.2004.01.008

Ben-Ami Bartal, I., Decety, J., and Mason, P. (2011). Empathy and pro-social behavior in rats. *Science* 334, 1427–1430. doi: 10.1126/science.1210789

Ben-Ami Bartal, I., Rodgers, D. A., Bernardez Sarria, M. S., Decety, J., and Mason, P. (2014). Pro-social behavior in rats is modulated by social experience. *eLife* 3:e01385. doi: 10.7554/eLife.01385

Ben-Ami Bartal, I., Shan, H., Molasky, N. M., Murray, T. M., Williams, J. Z., Decety, J., et al. (2016). Anxiolytic treatment impairs helping behavior in rats. *Front. Psychol.* 7:850. doi: 10.3389/fpsyg.2016.00850

Blake, R., and Shiffrar, M. (2007). Perception of human motion. *Annu. Rev. Psychol.* 58, 47–73. doi: 10.1146/annurev.psych.57.102904.190152

Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., et al. (2011). Early specialization for voice and emotion processing in the infant brain. *Curr. Biol.* 21, 1220–1224. doi: 10.1016/j.cub.2011.06.009

Boyatzis, C. J., Chazan, E., and Ting, C. Z. (1993). Preschool children's decoding of facial emotions. *J. Genet. Psychol.* 154, 375–382. doi: 10.1080/00221325.1993.10532190

Brosnan, S. F. (2013). Justice- and fairness-related behaviors in nonhuman primates. *Proc. Natl. Acad. Sci. U.S.A.* 110(Suppl. 2), 10416–10423. doi: 10.1073/pnas.1301194110

Carver, L. J., Dawson, G., Panagiotides, H., Meltzoff, A. N., McPartland, J., Gray, J., et al. (2003). Age-related differences in neural correlates of face recognition during the toddler and preschool years. *Dev. Psychobiol.* 42, 148–159. doi: 10.1002/dev.10078

Chen, C., Chen, C. Y., Yang, C. Y., Lin, C. H., and Cheng, Y. (2015). Testosterone modulates preattentive sensory processing and involuntary attention switches to emotional voices. *J. Neurophysiol.* 113, 1842–1849. doi: 10.1152/jn.00587.2014

Chen, C., Hu, C. H., and Cheng, Y. (2017a). Mismatch negativity (MMN) stands at the crossroads between explicit and implicit emotional processing. *Hum. Brain Mapp.* 38, 140–150. doi: 10.1002/hbm.23349

Chen, C., Hung, A. Y., Fan, Y. T., Tan, S., Hong, H., and Cheng, Y. (2017b). Linkage between pain sensitivity and empathic response in adolescents with autism spectrum conditions and conduct disorder symptoms. *Autism Res.* 10, 267–275. doi: 10.1002/aur.1653

Chen, C., Lee, Y. H., and Cheng, Y. (2014). Anterior insular cortex activity to emotional salience of voices in a passive oddball paradigm. *Front. Hum. Neurosci.* 8:743. doi: 10.3389/fnhum.2014.00743

Chen, C., Liu, C. C., Weng, P. Y., and Cheng, Y. (2016a). Mismatch negativity to threatening voices associated with positive symptoms in schizophrenia. *Front. Hum. Neurosci.* 10:362. doi: 10.3389/fnhum.2016.00362

Chen, C., Sung, J. Y., and Cheng, Y. (2016b). Neural dynamics of emotional salience processing in response to voices during the stages of sleep. *Front. Behav. Neurosci.* 10:117. doi: 10.3389/fnbeh.2016.00117

Chen, C., Yang, C. Y., and Cheng, Y. (2012). Sensorimotor resonance is an outcome but not a platform to anticipating harm to others. *Soc. Neurosci.* 7, 578–590. doi: 10.1080/17470919.2012.686924

Cheng, Y., Chen, C., and Decety, J. (2014). An EEG/ERP investigation of the development of empathy in early and middle childhood. *Dev. Cogn. Neurosci.* 10, 160–169. doi: 10.1016/j.dcn.2014.08.012

Cheng, Y., Hung, A. Y., and Decety, J. (2012a). Dissociation between affective sharing and emotion understanding in juvenile psychopaths. *Dev. Psychopathol.* 24, 623–636. doi: 10.1017/S095457941200020X

Cheng, Y., Lee, S. Y., Chen, H. Y., Wang, P. Y., and Decety, J. (2012b). Voice and emotion processing in the human neonatal brain. *J. Cogn. Neurosci.* 24, 1411–1419. doi: 10.1162/jocn_a_00214

Cheour, M., Alho, K., Ceponiene, R., Reinikainen, K., Sainio, K., Pohjavuori, M., et al. (1998). Maturation of mismatch negativity in infants. *Int. J. Psychophysiol.* 29, 217–226. doi: 10.1016/S0167-8760(98)00017-8

Church, R. M. (1959). Emotional reactions of rats to the pain of others. *J. Comp. Physiol. Psychol.* 52, 132–134. doi: 10.1037/h0043531

Cowell, J. M., and Decety, J. (2015a). Precursors to morality in development as a complex interplay between neural, socio environmental, and behavioral facets. *Proc. Natl. Acad. Sci. U.S.A.* 112, 12657–12662. doi: 10.1073/pnas.1508832112

Cowell, J. M., and Decety, J. (2015b). The neuroscience of implicit moral evaluation and its relation to generosity in early childhood. *Curr. Biol.* 25, 93–97. doi: 10.1016/j.cub.2014.11.002

de Waal, F. B. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annu. Rev. Psychol.* 59, 279–300. doi: 10.1146/annurev.psych.59.103006.093625

Decety, J. (2010). The neurodevelopment of empathy in humans. *Dev. Neurosci.* 32, 257–267. doi: 10.1159/000317771

Decety, J., Bartal, I. B., Uzefovsky, F., and Knafo-Noam, A. (2016). Empathy as a driver of prosocial behaviour: highly conserved neurobehavioural mechanisms across species. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:20150077. doi: 10.1098/rstb.2015.0077

Decety, J., and Cowell, J. M. (2014a). Friends or foes: is empathy necessary for moral behavior? *Perspect. Psychol. Sci.* 9, 525–537. doi: 10.1177/1745691614545130

Decety, J., and Cowell, J. M. (2014b). The complex relation between morality and empathy. *Trends Cogn. Sci.* 18, 337–339. doi: 10.1016/j.tics.2014.04.008

Decety, J., and Cowell, J. M. (2018). Interpersonal harm aversion as a necessary foundation for morality: a developmental neuroscience perspective. *Dev. Psychopathol.* 30, 153–164. doi: 10.1017/S0954579417000530

Decety, J., Meidenbauer, K. L., and Cowell, J. M. (2018). The development of cognitive empathy and concern in preschool children: a behavioral neuroscience investigation. *Dev. Sci.* 21:e12570. doi: 10.1111/desc.12570

Decety, J., and Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Dev. Cogn. Neurosci.* 2, 1–24. doi: 10.1016/j.dcn.2011.05.003

Decety, J., and Yoder, K. J. (2017). The emerging social neuroscience of justice motivation. *Trends Cogn. Sci.* 21, 6–14. doi: 10.1016/j.tics.2016.10.008

Dehaene-Lambertz, G., and Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature* 370, 292–295. doi: 10.1038/370292a0

Dehaene-Lambertz, G., and Pena, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *Neuroreport* 12, 3155–3158. doi: 10.1097/00001756-200110080-00034

Dinstein, I., Thomas, C., Behrmann, M., and Heeger, D. J. (2008). A mirror up to nature. *Curr. Biol.* 18, R13–R18. doi: 10.1016/j.cub.2007.11.004

Dondi, M., Simion, F., and Caltran, G. (1999). Can newborns discriminate between their own cry and the cry of another newborn infant? *Dev. Psychol.* 35, 418–426.

Downing, P. E., Jiang, Y., Shuman, M., and Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science* 293, 2470–2473. doi: 10.1126/science.1063414

Fan, Y. T., Chen, C., Chen, S. C., Decety, J., and Cheng, Y. (2014). Empathic arousal and social understanding in individuals with autism: evidence from fMRI and ERP measurements. *Soc. Cogn. Affect. Neurosci.* 9, 1203–1213. doi: 10.1093/scan/nst101

Fan, Y. T., and Cheng, Y. (2014). Atypical mismatch negativity in response to emotional voices in people with autism spectrum conditions. *PLoS One* 9:e102471. doi: 10.1371/journal.pone.0102471

Fan, Y. T., Hsu, Y. Y., and Cheng, Y. (2013). Sex matters: n-back modulates emotional mismatch negativity. *Neuroreport* 24, 457–463. doi: 10.1097/WNR.0b013e32836169b9

Ferguson, H. J., Cane, J. E., Douchkov, M., and Wright, D. (2015). Empathy predicts false belief reasoning ability: evidence from the N400. *Soc. Cogn. Affect. Neurosci.* 10, 848–855. doi: 10.1093/scan/nsu131

Frith, C. D., and Frith, U. (2006). The neural basis of mentalizing. *Neuron* 50, 531–534. doi: 10.1016/j.neuron.2006.05.001

Hamann, K., Warneken, F., Greenberg, J. R., and Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature* 476, 328–331. doi: 10.1038/nature10278

Hamlin, J. K. (2014). The case for social evaluation in preverbal infants: gazing toward one's goal drives infants' preferences for Helpers over Hinderers in the hill paradigm. *Front. Psychol.* 5:1563. doi: 10.3389/fpsyg.2014.01563

Hamlin, J. K. (2015). "The infantile origins of our moral brains," in *The Moral Brain: A Multidisciplinary Perspective*, eds J. Decety and T. Wheatley (Cambridge, MA: MIT Press), 105–122.

Hamlin, J. K., Wynn, K., and Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Dev. Sci.* 13, 923–929. doi: 10.1111/j.1467-7687.2010.00951.x

Hatfield, E., Cacioppo, J. T., and Rapson, R. L. (1993). Emotional contagion. *Curr. Dir. Psychol. Sci.* 2, 96–99. doi: 10.1111/1467-8721.ep10770953

Haviland, J. M., and Lelwica, M. (1987). The induced affect response: 10-week-old infants' responses to three emotion expressions. *Dev. Psychol.* 23, 97–104. doi: 10.1037/0012-1649.23.1.97

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* 21, 1229–1243. doi: 10.1162/jocn.2009.21189

Hoffman, M. L. (2007). "The origins of empathic morality in toddlerhood," in *Socioemotional Development in the Toddler Years: Transitions and Transformations*, eds C. A. Brownell and C. B. Kopp (New York, NY: Guilford Press), 132–145.

Hung, A. Y., Ahveninen, J., and Cheng, Y. (2013). Atypical mismatch negativity to distressful voices associated with conduct disorder symptoms. *J. Child Psychol. Psychiatry* 54, 1016–1027. doi: 10.1111/jcpp.12076

Hung, A. Y., and Cheng, Y. (2014). Sex differences in preattentive perception of emotional voices and acoustic attributes. *Neuroreport* 25, 464–469. doi: 10.1097/WNR.0000000000000115

Kanakogi, Y., Okumura, Y., Inoue, Y., Kitazaki, M., and Itakura, S. (2013). Rudimentary sympathy in preverbal infants: preference for others in distress. *PLoS One* 8:e65292. doi: 10.1371/journal.pone.0065292

Kilner, J. M., Friston, K. J., and Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cogn. Process.* 8, 159–166. doi: 10.1007/s10339-007-0170-2

Kisilevsky, B. S., Hains, S. M., Lee, K., Xie, X., Huang, H., Ye, H. H., et al. (2003). Effects of experience on fetal voice recognition. *Psychol. Sci.* 14, 220–224. doi: 10.1111/1467-9280.02435

Kret, M. E., Pichon, S., Grezes, J., and de Gelder, B. (2011). Similarities and differences in perceiving threat from dynamic faces and bodies. An fMRI study. *Neuroimage* 54, 1755–1762. doi: 10.1016/j.neuroimage.2010.08.012

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843. doi: 10.1038/nrn1533

Lee, Y. E., Yun, J. E., Kim, E. Y., and Song, H. J. (2015). The development of infants' sensitivity to behavioral intentions when inferring others' social preferences. *PLoS One* 10:e0135588. doi: 10.1371/journal.pone.0135588

Lepage, J. F., and Theoret, H. (2007). The mirror neuron system: grasping others' actions from birth? *Dev. Sci.* 10, 513–523. doi: 10.1111/j.1467-7687.2007.00631.x

Li, W., and Han, S. (2010). Perspective taking modulates event-related potentials to perceived pain. *Neurosci. Lett.* 469, 328–332. doi: 10.1016/j.neulet.2009.12.021

Lloyd-Fox, S., Blasi, A., Mercure, E., Elwell, C. E., and Johnson, M. H. (2012). The emergence of cerebral specialization for the human voice over the first months of life. *Soc. Neurosci.* 7, 317–330. doi: 10.1080/17470919.2011.614696

Martin, G. B., and Clark, R. D. (1982). Distress crying in neonates: species and peer specificity. *Dev. Psychol.* 18, 3–9. doi: 10.1037/0012-1649.18.1.3

Mella, N., Studer, J., Gilet, A. L., and Labouvie-Vief, G. (2012). Empathy for pain from adolescence through adulthood: an event-related brain potential study. *Front. Psychol.* 3:501. doi: 10.3389/fpsyg.2012.00501

Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. *Dev. Sci.* 10, 126–134. doi: 10.1111/j.1467-7687.2007.00574.x

Michalska, K. J., Zeffiro, T. A., and Decety, J. (2016). Brain response to viewing others being harmed in children with conduct disorder symptoms. *J. Child Psychol. Psychiatry* 57, 510–519. doi: 10.1111/jcpp.12474

Nelson, N. L., and Mondloch, C. J. (2018). Children's visual attention to emotional expressions varies with stimulus movement. *J. Exp. Child Psychol.* 172, 13–24. doi: 10.1016/j.jecp.2018.03.001

Nystrom, P. (2008). The infant mirror neuron system studied with high density EEG. *Soc. Neurosci.* 3, 334–347. doi: 10.1080/17470910701563665

Oostenbroek, J., Suddendorf, T., Nielsen, M., Redshaw, J., Kennedy-Costantini, S., Davis, J., et al. (2016). Comprehensive longitudinal study challenges the existence of neonatal imitation in humans. *Curr. Biol.* 26, 1334–1338. doi: 10.1016/j.cub.2016.03.047

Peelen, M. V., Atkinson, A. P., Andersson, F., and Vuilleumier, P. (2007). Emotional modulation of body-selective visual areas. *Soc. Cogn. Affect. Neurosci.* 2, 274–283. doi: 10.1093/scan/nsm023

Peelen, M. V., Atkinson, A. P., and Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *J. Neurosci.* 30, 10127–10134. doi: 10.1523/JNEUROSCI.2161-10.2010

Pichon, S., de Gelder, B., and Grezes, J. (2008). Emotional modulation of visual and motor areas by dynamic body expressions of anger. *Soc. Neurosci.* 3, 199–212. doi: 10.1080/17470910701394368

Pichon, S., de Gelder, B., and Grezes, J. (2009). Two different faces of threat. Comparing the neural systems for recognizing fear and anger in dynamic body expressions. *Neuroimage* 47, 1873–1883. doi: 10.1016/j.neuroimage.2009.03.084

Price, S. A., and Brosnan, S. F. (2012). To each according to his need? Variability in the responses to inequity in non-human primates. *Soc. Justice Res.* 25, 140–169. doi: 10.1007/s11211-012-0153-z

Sagi, A., and Hoffman, M. L. (1976). Empathic distress in the newborn. *Dev. Psychol.* 12, 175–176. doi: 10.1037/0012-1649.12.2.175

Schirmer, A., Striano, T., and Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport* 16, 635–639. doi: 10.1097/00001756-200504250-00024

Schmidt, M. F., and Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS One* 6:e23223. doi: 10.1371/journal.pone.0023223

Scott, R. M., and Baillargeon, R. (2017). Early false-belief understanding. *Trends Cogn. Sci.* 21, 237–249. doi: 10.1016/j.tics.2017.01.012

Scott, R. M., and Roby, E. (2015). Processing demands impact 3-year-olds' performance in a spontaneous-response task: new evidence for the processing-load account of early false-belief understanding. *PLoS One* 10:e0142405. doi: 10.1371/journal.pone.0142405

Simner, M. L. (1971). Newborn's response to the cry of another infant. *Dev. Psychol.* 5, 136–150. doi: 10.1037/h0031066

Sinke, C. B., Sorger, B., Goebel, R., and de Gelder, B. (2010). Tease or threat? Judging social interactions from bodily expressions. *Neuroimage* 49, 1717–1727. doi: 10.1016/j.neuroimage.2009.09.065

Sloane, S., Baillargeon, R., and Premack, D. (2012). Do infants have a sense of fairness? *Psychol. Sci.* 23, 196–204. doi: 10.1177/0956797611422072

Sommerville, J. A., Schmidt, M. F. H., Yun, J. E., and Burns, M. (2013). The development of fairness expectations and prosocial behavior in the second year of life. *Infancy* 18, 40–66. doi: 10.1111/j.1532-7078.2012.00129.x

Tomasello, M., and Gonzalez-Cabrera, I. (2017). The role of ontogeny in the evolution of human cooperation. *Hum. Nat.* 28, 274–288. doi: 10.1007/s12110-017-9291-1

Tucker, D. M., Luu, P., and Derryberry, D. (2005). Love hurts: the evolution of empathic concern through the encephalization of nociceptive capacity. *Dev. Psychopathol.* 17, 699–713. doi: 10.1017/S0954579405050339

Vaish, A., Carpenter, M., and Tomasello, M. (2010). Young children selectively avoid helping people with harmful intentions. *Child Dev.* 81, 1661–1669. doi: 10.1111/j.1467-8624.2010.01500.x

Vaish, A., Missana, M., and Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions. *Br. J. Dev. Psychol.* 29(Pt 1), 124–130. doi: 10.1348/026151010X532888

Warneken, F., Grafenhain, M., and Tomasello, M. (2012). Collaborative partner or social tool? New evidence for young children's understanding of joint intentions in collaborative activities. *Dev. Sci.* 15, 54–61. doi: 10.1111/j.1467-7687.2011.01107.x

Warneken, F., and Tomasello, M. (2009a). The roots of human altruism. *Br. J. Psychol.* 100(Pt 3), 455–471. doi: 10.1348/000712608X379061

Warneken, F., and Tomasello, M. (2009b). Varieties of altruism in children and chimpanzees. *Trends Cogn. Sci.* 13, 397–402. doi: 10.1016/j.tics.2009.06.008

Woodward, A. L., and Gerson, S. A. (2014). Mirroring and the development of action understanding. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130181. doi: 10.1098/rstb.2013.0181

Zhang, D., Liu, Y., Hou, X., Sun, G., Cheng, Y., and Luo, Y. (2014). Discrimination of fearful and angry emotional voices in sleeping human neonates: a study of the mismatch brain responses. *Front. Behav. Neurosci.* 8:422. doi: 10.3389/fnbeh.2014.00422

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

World Health Organization

Media centre

Up to 650 000 people die of respiratory diseases linked to seasonal flu each year

News release
13 December 2017

Up to 650 000 deaths annually are associated with respiratory diseases from seasonal influenza, according to new estimates by the United States Centers for Disease Control and Prevention (US-CDC), WHO and global health partners.

This marks an increase on the previous global estimate of 250 000-500 000, which dates from over ten years ago and covered all influenza-related deaths, including cardiovascular disease or diabetes. The new figures of 290 000-650 000 deaths are based on more recent data from a larger, more diverse group of countries, including lower middle-income countries, and exclude deaths from non-respiratory diseases.

“These figures indicate the high burden of influenza and its substantial social and economic cost to the world,” said Dr Peter Salama, Executive Director of WHO’s Health Emergencies Programme. “They highlight the importance of influenza prevention for seasonal epidemics, as well as preparedness for pandemics.”

The estimates take into account findings from recent influenza respiratory mortality studies, including a study conducted by the United States Centers for Disease Control and Prevention (US-CDC), published in The Lancet on Thursday (14 December).

According to US-CDC, most deaths occur among people aged over 75 years, and in the world’s poorest regions. Sub-Saharan Africa accounts for the world’s greatest flu mortality risk, followed closely by the Eastern Mediterranean and Southeast Asia.

“All countries, rich and poor, large and small, must work together to control influenza outbreaks before the arrival of the next pandemic. This includes building capacity to detect and respond to outbreaks, and strengthening health systems to improve the health of the most vulnerable and those most at risk,” said Dr Salama.

Nearly all deaths among children under 5 with influenza-related lower respiratory tract infections occur in developing countries, but the effects

of seasonal influenza epidemics on the world's poorest are not fully known.

WHO is working with partners to assess the global influenza burden of disease by providing guidance and expertise to Member States to measure the influenza disease burden and its economic consequences.

Further surveillance and laboratory studies of other diseases such as cardiovascular disease, which can be influenza-related, are expected to yield substantially higher estimates over the next few years.

WHO encourages countries to prioritize influenza prevention and produce national estimates to inform prevention policies. Annual influenza vaccination is recommended to prevent disease and complications from influenza infection. Vaccination is especially important for people at higher risk of serious influenza complications and death, and for health workers.

Seasonal influenza is an acute viral infection that spreads easily from person to person and circulates worldwide. Most people recover within a week without requiring medical attention. Common respiratory diseases related to seasonal influenza that can cause death include pneumonia and bronchitis.

WHO's Influenza Burden of Disease Working Group comprises experts from the All India Institute of Medical Science, the National University of Singapore, the South African National Institute of Communicable Diseases, US CDC, Universidad del Valle de Guatemala and the University of Edinburgh.

Influenza fact sheet

Media contact:

Gregory HÄRTL
Coordinator, News, Social Media and Monitoring
Telephone : +41 22 791 4458
Mobile: +41 79 203 6715
Email: hartlg@who.int

What we do

Regions

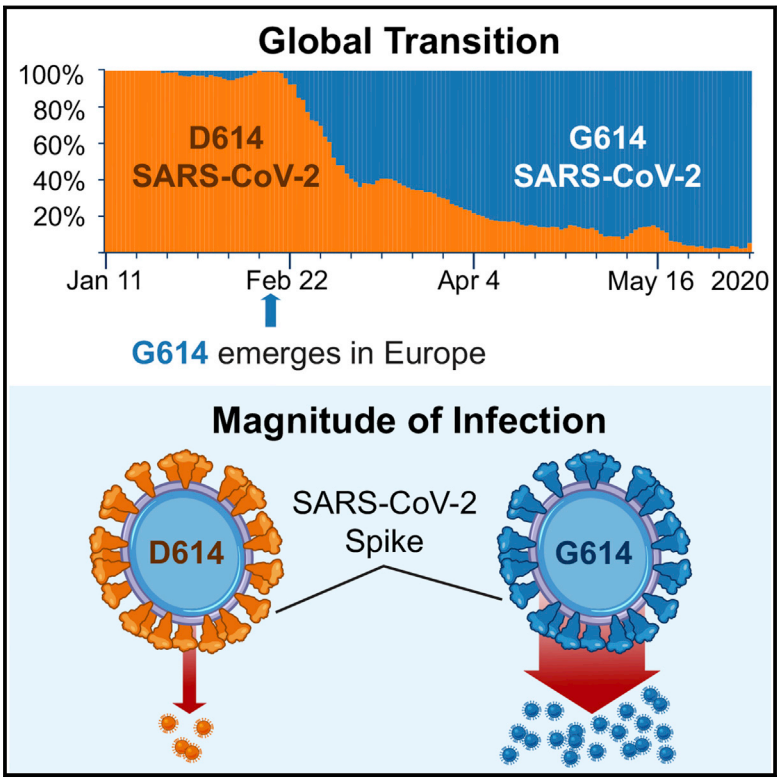
About us

Subscribe to our newsletters

Cell

Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus

Graphical Abstract



Highlights

- A SARS-CoV-2 variant with Spike G614 has replaced D614 as the dominant pandemic form
- The consistent increase of G614 at regional levels may indicate a fitness advantage
- G614 is associated with lower RT PCR Cts, suggestive of higher viral loads in patients
- The G614 variant grows to higher titers as pseudotyped virions

Authors

Bette Korber, Will M. Fischer, Sandrasegaram Gnanakaran, ..., Celia C. LaBranche, Erica O. Saphire, David C. Montefiori

Correspondence

btk@lanl.gov

In Brief

Korber et al. present evidence that there are now more SARS-CoV-2 viruses circulating in the human population globally that have the G614 form of the Spike protein versus the D614 form that was originally identified from the first human cases in Wuhan, China. Follow-up studies show that patients infected with G614 shed more viral nucleic acid compared with those with D614, and G614-bearing viruses show significantly higher infectious titers *in vitro* than their D614 counterparts.

Korber et al., 2020, Cell *182*, 812–827
August 20, 2020 Published by Elsevier Inc.
https://doi.org/10.1016/j.cell.2020.06.043

CellPress

CellPress
OPEN ACCESS

Cell

Article

Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus

Bette Korber,[1,2,10,*] Will M. Fischer,[1] Sandrasegaram Gnanakaran,[1] Hyejin Yoon,[1] James Theiler,[1] Werner Abfalterer,[1] Nick Hengartner,[1] Elena E. Giorgi,[1] Tanmoy Bhattacharya,[1] Brian Foley,[1] Kathryn M. Hastie,[3] Matthew D. Parker,[4] David G. Partridge,[5] Cariad M. Evans,[5] Timothy M. Freeman,[4] Thushan I. de Silva,[5,6] on behalf of the Sheffield COVID-19 Genomics Group, Charlene McDanal,[7] Lautaro G. Perez,[7] Haili Tang,[7] Alex Moon-Walker,[3,8,9] Sean P. Whelan,[9] Celia C. LaBranche,[7] Erica O. Saphire,[3] and David C. Montefiori[7]

[1]Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[2]New Mexico Consortium, Los Alamos, NM 87545, USA
[3]La Jolla Institute for Immunology, La Jolla, CA 92037, USA
[4]Sheffield Biomedical Research Centre & Sheffield Bioinformatics Core, University of Sheffield, Sheffield S10 2HQ, UK
[5]Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield S10 2JF, UK
[6]Department of Infection, Immunity and Cardiovascular Disease, Medical School, University of Sheffield, Sheffield S10 2RX, UK
[7]Duke Human Vaccine Institute & Department of Surgery, Durham, NC 27710, USA
[8]Program in Virology, Harvard University, Boston, MA 02115, USA
[9]Department of Molecular Microbiology, Washington University in Saint Louis, St. Louis, MO 63130, USA
[10]Lead Contact
*Correspondence: btk@lanl.gov
https://doi.org/10.1016/j.cell.2020.06.043

SUMMARY

A SARS-CoV-2 variant carrying the Spike protein amino acid change D614G has become the most prevalent form in the global pandemic. Dynamic tracking of variant frequencies revealed a recurrent pattern of G614 increase at multiple geographic levels: national, regional, and municipal. The shift occurred even in local epidemics where the original D614 form was well established prior to introduction of the G614 variant. The consistency of this pattern was highly statistically significant, suggesting that the G614 variant may have a fitness advantage. We found that the G614 variant grows to a higher titer as pseudotyped virions. In infected individuals, G614 is associated with lower RT-PCR cycle thresholds, suggestive of higher upper respiratory tract viral loads, but not with increased disease severity. These findings illuminate changes important for a mechanistic understanding of the virus and support continuing surveillance of Spike mutations to aid with development of immunological interventions.

INTRODUCTION

The past two decades have seen three major pathogenic zoonotic disease outbreaks caused by betacoronaviruses (Cui et al., 2019; de Wit et al., 2016; Liu et al., 2020; Wu et al., 2020). Severe acute respiratory syndrome coronavirus (SARS-CoV) emerged in 2002, infecting ~8,000 people with a 10% mortality. Middle East respiratory syndrome coronavirus (MERS-CoV) emerged in 2012 with ~2,300 cases and 35% mortality (Graham and Baric, 2010). The third, SARS-CoV-2, causes the severe respiratory disease coronavirus disease 2019 (COVID-19) (Gorbalenya et al., 2020). First reported in China in December 2019 (Zhou et al., 2020), it rapidly became a pandemic with devastating effects. The June 21, 2020 World Health Organization (WHO) Situation Report records over 8.7 million COVID-19 cases and 460,000 deaths, numbers that increase daily. Humans have no direct immunological experience with SARS-CoV-2, leaving us vulnerable to infection and disease. SARS-CoV-2 is highly transmissible: basic reproduction number, $R_{0,}$estimates vary between 2.2 and 3.9 (Lv et al., 2020). Estimates of mortality vary regionally between 0.8% and 14.5% (mortality analyses, Johns Hopkins University of Medicine)

Coronaviruses have genetic proofreading mechanisms (Sevajol et al., 2014; Smith et al., 2013), and SARS-CoV-2 sequence diversity is very low (Fauver et al., 2020). Still, natural selection can act upon rare but favorable mutations. By analogy, antigenic drift results in gradual accumulation of mutations by the influenza virus during flu season, and the complex interplay between immunological resistance mutations and the fitness landscape enables antibody resistance to develop across populations (Wu et al., 2020), driving the need to develop new influenza vaccines every few seasons. Longer flu seasons have increased opportunities for selection pressure (Boni et al., 2006). Although SARS-CoV-2 shows evidence of some seasonal waning (Sehra

et al., 2020), the persistence of the pandemic may enable accumulation of immunologically relevant mutations in the population even as vaccines are developed. Antigenic drift is seen among the common cold coronaviruses OC43 (Ren et al., 2015; Vijgen et al., 2005) and 229E (Chibo and Birch, 2006) and in SARS-CoV-1 (Guan et al., 2003; Song et al., 2005). Notably, a single SARS-CoV-1 amino acid change, Spike D480A/G in the receptor binding domain (RBD), arose in infected humans and civets and became the dominant variant among 2003/2004 viruses. D480A/G escapes neutralizing antibody 80R, and immune pressure from 80R *in vitro* could recapitulate emergence of the D480 mutation (Sui et al., 2008). Although there is no evidence yet of antigenic drift for SARS-CoV-2, with extended human-to-human transmission, SARS-CoV-2 could also acquire mutations with fitness advantages and immunological resistance. Attending to this risk now by identifying evolutionary transitions that may be relevant to the fitness or antigenic profile of the virus is important to ensure effectiveness of the vaccines and immunotherapeutic interventions as they advance to the clinic.

In response to the urgent need to develop effective vaccines and antibody-based therapeutic agents against SARS-CoV-2, over 90 vaccine and 50 antibody approaches are currently being explored (Cohen, 2020; Yu et al., 2020). Most target the trimeric Spike protein, which mediates host cell binding and entry and is the major target of neutralizing antibodies (Chen et al., 2020; Yuan et al., 2020). Spike monomers are comprised of an N-terminal S_1 subunit that mediates receptor binding and a membrane-proximal S_2 subunit that mediates membrane fusion (Hoffmann et al., 2020a; Walls et al., 2020; Wrapp et al., 2020). SARS-CoV-2 and SARS-CoV-1 share ~79% sequence identity (Lu et al., 2020), and both use angiotensin-converting enzyme 2 (ACE2) as their cellular receptor. Antibody responses to SARS-CoV-1 Spike are complex. In some patients with rapid and high neutralizing antibody responses, an early decline of these responses is associated with increased severity of disease and a higher risk of death (Ho et al., 2005; Liu et al., 2006; Temperton et al., 2005; Zhang et al., 2006). Some antibodies against SARS-CoV-1 Spike mediate antibody-dependent enhancement (ADE) of infection *in vitro* and exacerbate disease in animal models (Jaume et al., 2011; Wan et al., 2020; Wang et al., 2014; Yip et al., 2016).

Most current SARS-CoV-2 immunogens and testing reagents are based on the Spike protein sequence of the Wuhan reference sequence (Wang et al., 2020), and first-generation antibody therapeutic agents were discovered based on early pandemic infections and evaluated using the Wuhan reference sequence proteins. Alterations of the reference sequence as the virus propagates in human-to-human transmission could potentially alter the viral phenotype and/or the efficacy of immune-based interventions. Therefore, we designed bioinformatics tools to create an "early warning" strategy to evaluate Spike evolution during the pandemic to enable testing of mutations for phenotypic implications and generation of appropriate antibody breadth evaluation panels as vaccines and antibody-based therapeutic agents progress. Phylogenetic analysis of the global sampling of SARS-CoV-2 is being very capably addressed by the Global Initiative for Sharing All Influenza Data (GISAID) database (https://www.gisaid.org/; Elbe and Buckland-Merrett, 2017; Shu and McCauley, 2017) and Nextstrain (https://nextstrain.org; Hadfield et al., 2018). However, in a setting of low genetic diversity like that of SARS-CoV-2, with very few *de novo* mutational events, phylogenetic methods that use homoplasy to identify positive selection (Crispell et al., 2019) have limited statistical power. Additionally, recombination can add a confounding factor to phylogenetic reconstructions, and recombination is known to play a role in natural coronavirus evolution (Graham and Baric, 2010; Lau et al., 2011; Li et al., 2020; Oong et al., 2017; Rehman et al., 2020), and recombinant sequences (potential sequencing artifacts) have been found among SARS-CoV-2 sequences (De Maio et al., 2020). Given these issues, we developed an alternative indicator of potential positive selection by identifying variants that are recurrently becoming more prevalent in different geographic locations. If increases in relative frequency of a particular variant are observed repeatedly in distinct geographic regions, then that variant becomes a candidate for conferring a selective advantage.

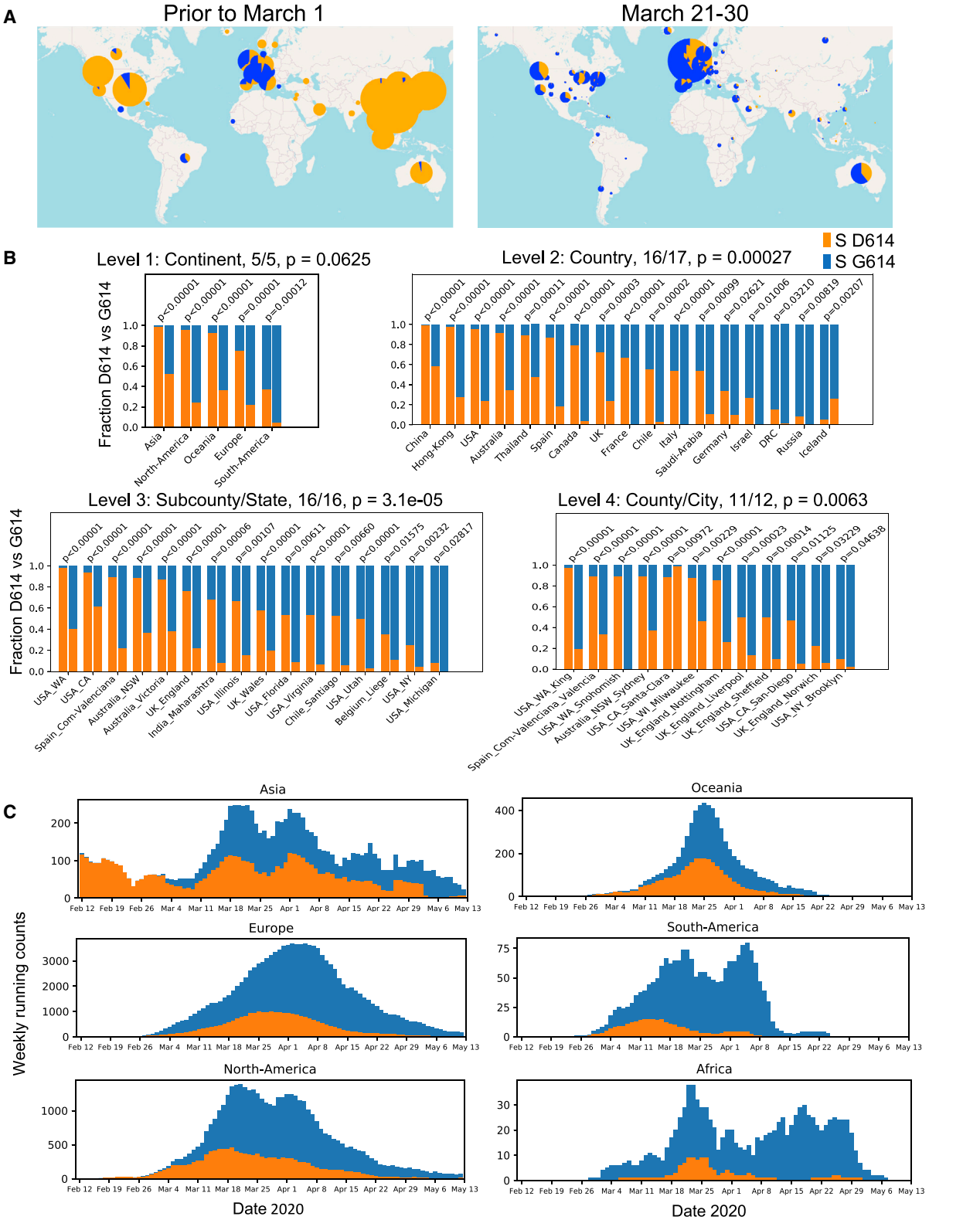
Single amino acid changes are worth monitoring because they can be phenotypically relevant. Among coronaviruses, point mutations have been demonstrated to confer resistance to neutralizing antibodies in MERS-CoV (Tang et al., 2014) and SARS-CoV-1 (Sui et al., 2008; ter Meulen et al., 2006). In the HIV envelope, single amino acid changes are known to alter host species susceptibility (Li et al., 2016), increase expression levels (Asmal et al., 2011), change the viral phenotype from tier 2 to tier 1, cause an overall change in neutralization sensitivity (Gao et al., 2014; LaBranche et al., 2019), and confer complete or nearly complete resistance to classes of neutralizing antibodies (Bricault et al., 2019; Sadjadpour et al., 2013; Zhou et al., 2019).

We developed a bioinformatics pipeline to identify Spike amino acid variants that are increasing in frequency across many geographic regions by monitoring GISAID data. By early April 2020, it was clear that the Spike D614G mutation exhibited this behavior, and G614 has since become the dominant form in the pandemic. We present experimental evidence that the G614 variant is associated with greater infectivity as well as clinical evidence that it is associated with higher viral loads. We continue to monitor other mutations in Spike for frequency shifts at regional and global levels and provide regular updates at a public web site (https://cov.lanl.gov/).

RESULTS

Website Overview

Our analysis pipeline to track SARS-CoV-2 mutations in the COVID-19 pandemic is based on regular updates from the GISAID SARS-CoV-2 sequence database (GISAID acknowledgments are in Table S1). GISAID sequences are generally linked to the location and date of sampling. Our website provides visualizations and summary data that allow regional tracking of SARS-CoV-2 mutations over time. Hundreds of new SARS-CoV-2 sequences are added to GISAID each day, so we have automated steps to create daily working alignments (Kurtz et al., 2004; Figure S1). The analysis presented here is based on a May 29, 2020 download of the GISAID data, when our Spike alignment included 28,576 sequences; updated versions of key figures can recreated at our website (https://cov.lanl.gov). The



(legend on next page)

overall evolutionary rate for SARS-CoV-2 is very low, so we set a low threshold for a Spike mutation to be deemed "of interest," and we track all sites in Spike where 0.3% of the sequences differ from the Wuhan reference sequence, monitoring them for increasing frequency over time in geographic regions as well as for recurrence in different geographic regions. Here we present results for the first amino acid variant to stand out by these metrics, D614G.

The D614G Variant

Increasing Frequency and Global Distribution

The Spike D614G amino acid change is caused by an A-to-G nucleotide mutation at position 23,403 in the Wuhan reference strain; it was the only site identified in our first Spike variation analysis in early March 2020 that met our threshold criterion. At that time, the G614 form was rare globally but gaining prominence in Europe, and GISAID was also tracking the clade carrying the D614G substitution, designating it the "G clade." The D614G change is almost always accompanied by three other mutations: a C-to-T mutation in the 5′ UTR (position 241 relative to the Wuhan reference sequence), a silent C-to-T mutation at position 3,037, and a C-to-T mutation at position 14,408 that results in an amino acid change in RNA-dependent RNA polymerase (RdRp P323L). The haplotype comprising these 4 genetically linked mutations is now the globally dominant form. Prior to March 1, 2020, it was found in 10% of 997 global sequences; between March 1 and March 31, 2020, it represented 67% of 14,951 sequences; and between April 1 and May 18, 2020 (the last data point available in our May 29, 2020 sample), it represented 78% of 12,194 sequences. The transition from D614 to G614 occurred asynchronously in different regions throughout the world, beginning in Europe, followed by North America and Oceania and then Asia (Figures 1, 2, 3, S2, and S3).

We developed two statistical approaches to assess the consistency and significance of the D614-to-G614 transition. In general, to observe a significant change in the frequency of variants in a geographic region, three requirements must be met. First, both variants must at some point be co-circulating in the geographic area. Second, there must be sampling over an adequate duration to observe a change in frequency. Third, enough samples must be available for adequate statistical power to detect a difference. Both of our approaches enable us to systematically extract all GISIAD local and regional data that meet these three requirements.

Our first approach requires that there be an "onset," defined as the first day where the cumulative number of sequences reached 15 and both forms were represented at least 3 times; we further require that there be at least 15 sequences available at least 2 weeks after onset. Each geographic region that meets these criteria is extracted separately based on the hierarchical geographic/political levels designated in GISAID (Figure 1B). A two-sided Fisher's exact test compares the counts in the pre-onset period with the counts after the 2-week delay period and provides a p value against the null hypothesis that the fraction of D614 versus G614 sequences did not change. All regions that met the above criteria and that showed a significant change in either direction ($p < 0.05$) are included. Almost all shifted toward increasing G614 frequencies: 5 of 5 continents, 16 of 17 countries (two-sided binomial p value of 0.00027), 16 of 16 regions ($p = 0.00003$), and 11 of 12 counties and cities ($p = 0.0063$).

In Figure 2 (Europe), Figure S2 (North America), and Figure S3 (Australia and Asia), we break down the relationships shown in Figure 1B in detail. The G614 variant increased in frequency even in regions where D614 was the clearly dominant form of a well-established local epidemic when G614 entered the population. Examples of this scenario include Wales, Nottingham, and Spain (Figure 2); Snohomish county and King county (Figure S2); and New South Wales, China, Japan, Hong Kong, and Thailand (Figure S3). Although introduction of a new variant might sometimes result in emergence of the new form because of stochastic effects or serial re-introductions or apparent emergence because of sampling biases, the consistency of the shift to G614 across regions is striking. The increase in G614 often continued after national stay-at-home orders were implemented and, in some cases, beyond the 2-week maximum incubation period.

We found two exceptions to the pattern of increasing G614 frequency in Figure 1B; details regarding these cases are shown in Figure S4. The first is Iceland. Changes in sampling strategy during a regional molecular epidemiology survey conducted through the month of March 2020 might explain this exception (Gudbjartsson et al., 2020). In early March 2020, only high-risk people were sampled, the majority being travelers from countries in Europe where G614 dominated. In mid-March 2020, screening began to include the local population; this coincided with the appearance of the D614 variant in the sequence dataset. The second exception is Santa Clara county, one of the most heavily sampled regions in California (Deng et al., 2020). The D614 variant dominates sequences from the Santa Clara

Figure 1. The Global Transition from the Original D614 Form to the G614 Variant

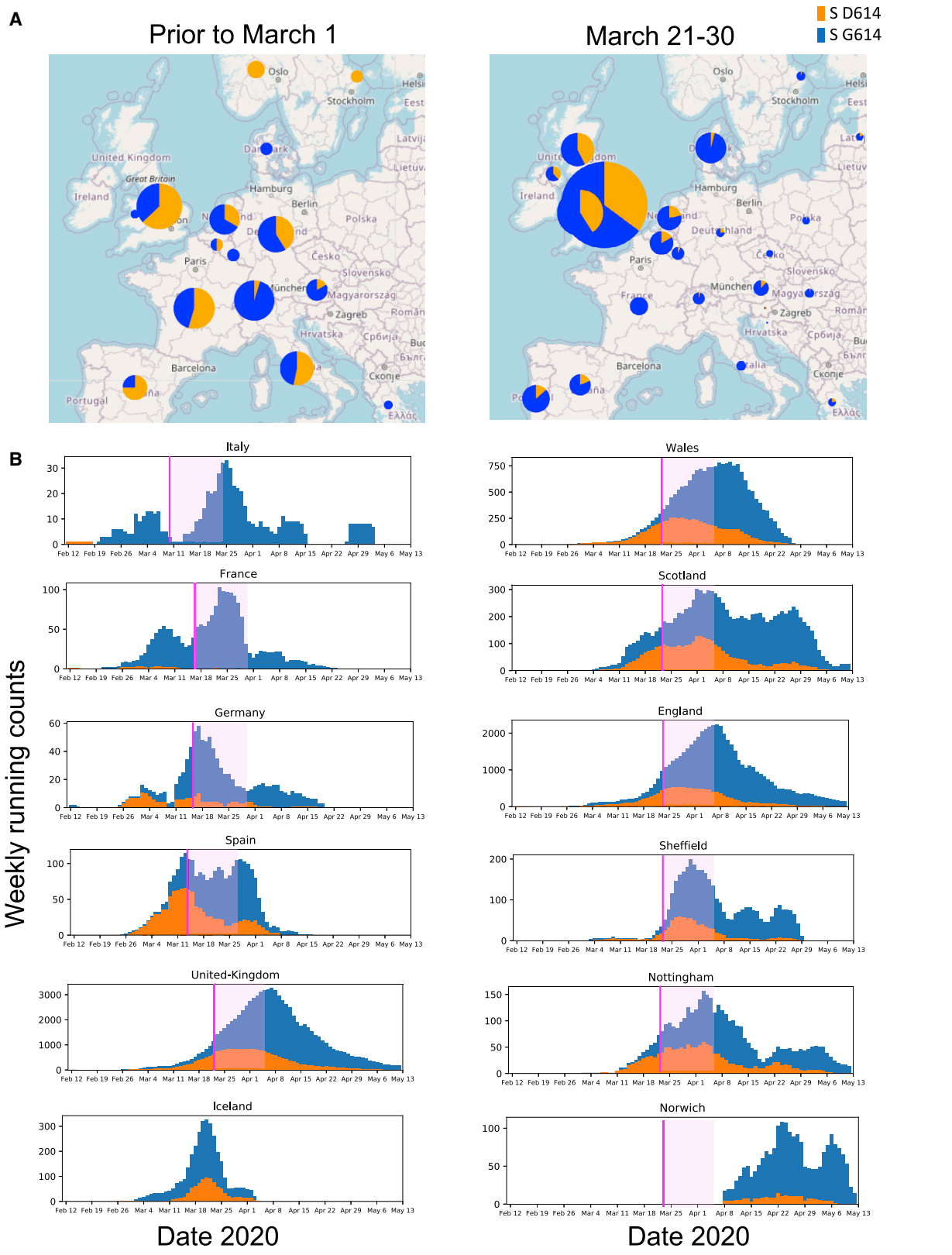
Figure360▷ For a Figure360 author presentation of this figure, see https://doi.org/10.1016/j.cell.2020.06.043#mmc4.

(A) Changes in the global distribution of the relative frequencies of the D614 (orange) and G614 (blue) variants in 2 time frames. Circle size indicates the relative sampling within each map. Through March 1, 2020, the G614 variant was rare outside of Europe, but by the end of March 2020 it had increased in frequency worldwide. These data are explored regionally in Figure 2 (Europe), Figure S2 (North America), and Figure S3 (Australia and Asia).

(B) Paired bar charts compare the fraction of sequences with D614 and with G614 for two time periods separated by a 2-week gap. The first time period (left bar) includes all sequences up to the onset day (see main text). The second time period (right bar) includes all sequences acquired at least 2 weeks after the onset date. All regions are shown that met the minimal threshold criteria for inclusion (see main text) with a significant shift in frequency (two-sided Fisher's exact test, $p < 0.05$). Four hierarchical geographic levels are split out based on GISAID naming conventions.

(C) Running weekly average counts of sampled sequences exhibiting the D614 (orange) and G614 (blue) variants on different continents between January 12 and May 12, 2020. The measure of interest is the relative frequency over time. The shape of the overall curve just reflects sample availability; sequencing was more limited earlier in the epidemic (hence the left-hand tail), and there is a time lag between viral sampling and sequence availability in GISAID (hence the right-hand tail).

Weekly running count plots were generated with Python Matplotlib (Hunter, 2007); all elements of this figure are frequently updated at https://cov.lanl.gov/.



(legend on next page)

Department of Public Health (DPH) to date; the G614 variant was apparently not established in that community. In contrast, a smaller set of Santa Clara county sequences, sampled from mid-March to early April 2020, were specifically noted to be from Stanford; the Stanford samples had a mixture of both forms co-circulating (Figure S4), suggesting that the two communities in Santa Clara County are effectively distinct. A June 19, 2020 GISAID update for several California counties is provided in Figure S4C, and the G614 form is present in the most recent Santa Clara DPH samples.

Our second statistical approach to evaluating the significance of the D614-to-G614 transition (Figure 3) uses the time series data in GISAID more fully. Here we extracted all regional data from GISAID that had a minimum of 5 sequences representing each of the D614 and the G614 variants and at least 14 days of sampling. We then modeled the daily fraction of G614 as a function of time using isotonic regression, testing the null hypothesis that this fraction does not change over time (i.e., it remains roughly flat over time with equally likely random fluctuations of increase or decrease). We then separately tested the null against two alternative hypotheses: that the fraction of G614 increases or that it decreases. Figure 3A shows separate p values for all subcountries/states and counties/cities that met the minimal criteria. 30 of 31 subcountries/states with a significant change in frequency were increasing in G614; a binomial test indicates that G614 increases are highly significantly enriched ($p = 2.98e{-}09$). This was also found in 17 of 19 counties/cites ($p = 0.0007$). Figure 3B shows examples for 3 cities, plotting the daily fraction of G614 as a function of time. Country summaries (similar to Figure 3A) and plots for all regions (similar to Figure 3B) are included in Data S1.

Origins of the D614G 4-Base Haplotype

The earliest examples of sequences carrying parts of the 4-mutation haplotype that characterizes the D614G GISAID G clade were found in China and Germany in late January 2020, and they carried 3 of the 4 mutations that define the clade, lacking only the RdRp P323L substitution (Figure S5D). This may be an ancestral form of the G clade. One early Wuhan sequence and one early Thai sequence had the D614G change but not the other 3 mutations (Figure S5D); these may have arisen independently. The earliest sequence we detected that carried all 4 mutations was sampled in Italy on February 20, 2020 (Figure S5D). Within days, this haplotype was sampled in many countries in Europe.

Structural Implications of the Spike D614G Change

D614 is located on the surface of the Spike protein protomer, where it can form contacts with the neighboring protomer (Figure 4A). Cryoelectron microscopy (cryo-EM) structures (Walls et al., 2020; Wrapp et al., 2020) indicate that the side chains of D614 and T859 of the neighboring protomer (Figure 4B) form a between-protomer hydrogen bond, bringing together a residue from the S_1 unit of one protomer and a residue of the S_2 unit of the other protomer (Figure 4C). The change to G614 would eliminate this side-chain hydrogen bond, possibly increasing main-chain flexibility and altering between-protomer interactions. In addition, this substitution could modulate glycosylation at the nearby N616 site, influence the dynamics of the spatially proximal fusion peptide (Figure 4D) of the neighboring protomer, or have other effects.

G614 Is Associated with Potentially Higher Viral Loads in COVID-19 Patients but Not with Disease Severity

SARS-CoV-2 sequences from 999 individuals presenting with COVID-19 disease at the Sheffield Teaching Hospitals NHS Foundation Trust were available and linked to clinical data. The Sheffield data include age, sex, date of sampling, hospitalization status (defined as outpatient [OP], inpatient [IP], requiring hospitalization, or admittance to the intensive care unit [ICU]), and the cycle threshold (Ct) for a positive signal in E-gene based RT-PCR. The Ct is used here as a surrogate for relative viral loads; lower Ct values indicate higher viral loads (Corman et al., 2020), but not all viral nucleic acids represent infectious viral particles. RT-PCR methods changed during the course of the study because of limited availability of testing kits. The first method involved nucleic acid extraction and the second method heat treatment (Fomsgaard and Rosenstierne, 2020). A generalized linear model (GLM) used to predict the PCR Ct based on the RT-PCR method, sex, age, and D614G status showed only the RT-PCR method ($p < 2e{-}16$) and D614G status ($p = 0.037$) to be statistically significant (Figure 5A). Lower Ct values were observed in G614 infections. While our paper was in revision, G614-variant association with low Ct values *in vivo* (Figure 5) was reported independently by two other groups (Lorenzo-Redondo et al., 2020; Wagner et al., 2020) in preprints that have not yet been peer reviewed.

We found no significant association between D614G status and disease severity as measured by hospitalization outcomes. A comparison of D614G status and hospitalization (combining IP and ICU) was not significant ($p = 0.66$, Fisher's exact test), although comparing ICU admission with IP and OP did have borderline significance ($p = 0.047$) (Figure 5B). Regression analysis reinforced the result that G614 status was not associated with greater levels of hospitalization but that higher age (Dowd et al., 2020; Promislow, 2020), male sex (Conti and Younes,

Figure 2. The Transition from D614 to G614 in Europe

(A) Maps of relative D614 and G614 frequencies in Europe in 2 time frames.

(B) Weekly running counts of G614 illustrating the timing of its spread in Europe. The legend for Figure 1 explains how to read these figures. Some nations essentially had G614 epidemics when sampling began, but even in these cases, small traces of D614 found early were soon lost (e.g., France and Italy). The Italian epidemic started with the D614 clade, but Italy had the first sampled case of the full G614 haplotype and had shifted to all G614 samples prior to March 1, 2020 (Figure S5). European nations that began with a mixture of D614 and G614 most clearly reveal the frequency shifts (e.g., Germany, Spain, and the United Kingdom). The United Kingdom is richly sampled and so is subdivided into smaller regions (England, Wales, and Scotland) and then further divided to display two well-sampled English cities. Even in settings with very well-established D614 epidemics (e.g., Wales and Nottingham; see also Figures S2 and S3), G614 becomes prevalent soon after its appearance. The increase in G614 frequency often continues well after stay-at-home orders are in place (pink line) and past the subsequent 2-week incubation period (pink transparent box). The figures shown here can be recreated with contemporary data from GISAID at the http://cov.lanl.gov/ website. UK stay-at-home order dates were based on the date of the national proclamation, and others were documented on the web.

A

GISAID level 3: Subcounty/State

Of 31 regions with a clear direction, 30 are increasing: Binomial p = 2.98e-08

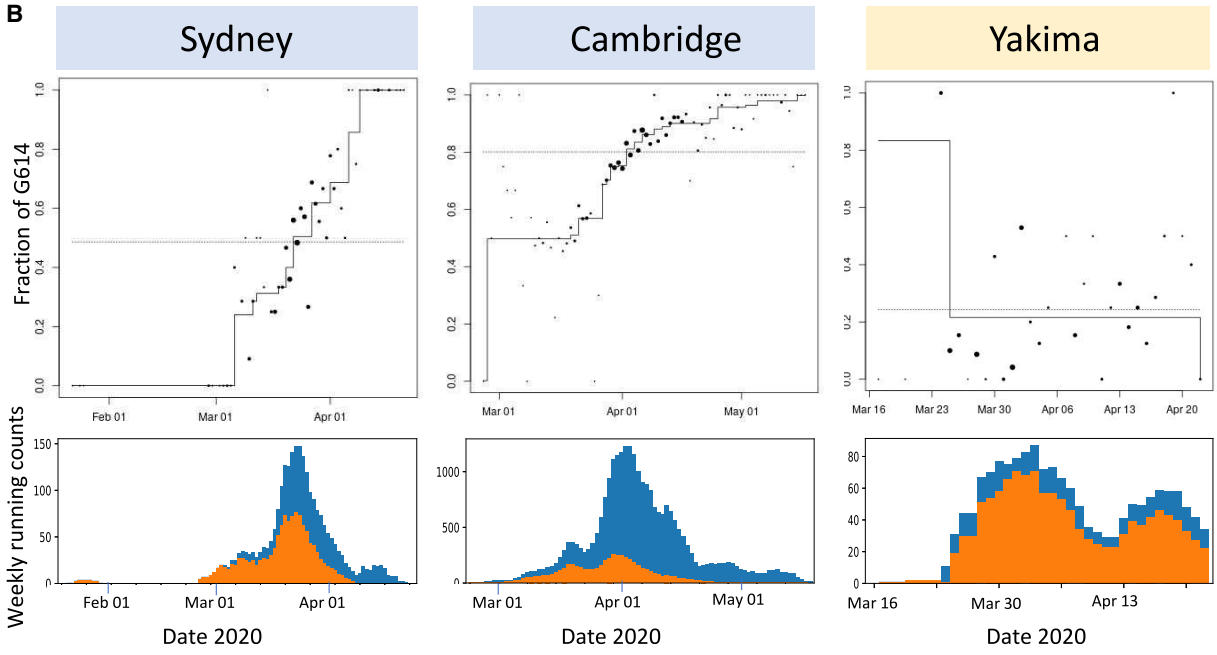
Level 2: Region	# D614	# G614	# of days	Time window days	G614 increasing p-value	G614 decreasing p-value
Australia_New-South-Wales	189	180	51	90	0.00025	0.99
Australia_Victoria	226	306	43	80	0.00025	0.93
Belgium_Ghent	12	37	18	26	0.00025	0.77
China_Beijing	46	25	35	66	0.00025	0.98
Germany_North-Rhine-Westphalia	16	37	16	27	0.00025	0.91
Spain_Comunitat-Valenciana	73	110	32	34	0.00025	0.95
Taiwan_Taoyuan	14	12	16	85	0.00025	0.51
United-Kingdom_England	1904	6338	88	109	0.00025	1.00
United-Kingdom_Scotland	433	1125	73	75	0.00025	0.68
United-Kingdom_Wales	717	1792	45	54	0.00025	0.95
USA_Arizona	11	64	22	71	0.00025	0.99
USA_California	267	199	70	111	0.00025	0.001*
*USA_California excluding Santa Clara	102	175	53	107	0.00025	0.99
USA_Michigan	31	382	38	53	0.00025	0.85
USA_New-York	91	1163	52	55	0.00025	1.00
USA_Utah	24	253	38	45	0.00025	0.98
USA_Virginia	27	220	43	47	0.00025	0.99
USA_Washington	926	683	69	101	0.00025	1.00
USA_Wisconsin	156	197	49	93	0.00025	0.07
India_Maharashtra	30	35	30	48	0.0005	0.97
Thailand_Bangkok	26	12	19	70	0.0005	0.72
Chile_Santiago	19	63	27	34	0.00075	0.99
USA_Illinois	38	56	26	88	0.00075	0.98
United-Kingdom_Northern-Ireland	47	60	22	27	0.00325	0.42
Australia_South-Australia	15	58	23	32	0.0035	0.84
USA_Minnesota	51	96	35	47	0.004	0.76
Taiwan_Taipei	17	26	26	94	0.0053	0.45
Australia_Queensland	12	13	14	58	0.0058	0.96
USA_Florida	11	35	18	69	0.009	0.82
USA_Connecticut	22	127	34	71	0.0095	0.93
Belgium_Liege	19	140	24	44	0.016	0.47
Denmark_Unknown	32	493	34	34	0.026	0.56
Canada_Ontario	26	32	15	57	0.062	0.91
India_Gujarat	10	110	22	36	0.085	0.08
Austria_Vienna	10	108	33	41	0.12	0.53
Spain_Madrid	20	56	16	33	0.15	0.89
Netherlands_Noord-Brabant	44	32	18	23	0.21	0.64
USA_Texas	75	169	29	53	0.26	0.72
Netherlands_Zuid-Holland	15	75	23	29	0.28	0.69
Netherlands_Utrecht	18	52	31	55	0.30	0.54
India_Delhi	24	13	16	32	0.93	0.00075

GISAID level 4: County/City

Of 19 cities with a clear direction, 17 are increasing: Binomial p = 0.0007

Level 3: County/City	# D614	# G614	# of days	Time window days	G614 increasing p-value	G614 decreasing p-value
Australia_New-South-Wales_Sydney	189	179	51	90	0.00025	1.00
Spain_Comunitat-Valenciana_Valencia	72	97	30	34	0.00025	0.64
United-Kingdom_England_Bristol	240	629	35	37	0.00025	0.28
United-Kingdom_England_Cambridge	751	3020	81	81	0.00025	1.00
United-Kingdom_England_Liverpool	97	484	46	45	0.00025	0.71
United-Kingdom_England_Nottingham	204	386	67	76	0.00025	0.99
United-Kingdom_England_Sheffield	120	431	44	51	0.00025	1.00
USA_Washington_King	173	75	58	69	0.00025	0.99
USA_Washington_Pierce	32	35	21	38	0.00025	1.00
USA_Washington_Snohomish	35	32	27	93	0.00025	1.00
USA_Wisconsin_Milwaukee	66	30	32	45	0.00025	0.97
United-Kingdom_England_Norwich	29	269	26	28	0.00075	0.97
USA_California_San-Diego	11	75	33	58	0.002	0.95
United-Kingdom_England_London	36	357	19	24	0.0085	0.91
USA_Wisconsin_Madison	13	43	26	35	0.030	0.39
USA_New-York_Manhattan	38	339	30	45	0.036	0.90
USA_California_San-Francisco	59	83	21	48	0.049	0.34
USA_New-York_Brooklyn	13	292	31	46	0.070	0.87
USA_Washington_Yakima	184	59	31	36	0.073	0.00025
USA_California_Santa-Clara	165	24	50	76	0.49	0.00025

B



(legend on next page)

2020; Promislow, 2020) and higher Ct values (lower viral loads) were highly predictive of hospitalization. Further analysis showed that viral load was not masking a potential D614G status effect on hospitalization (STAR Methods). Univariate analysis also found highly significant associations between age and male sex and hospitalization (STAR Methods).

G614 Is Associated with Higher Infectious Titers of Spike-Pseudotyped Virus

We quantified the infectious titers of pseudotyped single-cycle vesicular stomatitis virus (VSV) and lentiviral particles displaying D614 or G614 SARS-CoV2 Spike protein. For the VSV and lentiviral pseudotypes, G614-bearing viruses had significantly higher infectious titers (2.6- to 9.3-fold increase) than their D614 counterparts; this was confirmed in multiple cell types (Figures 6A–6C). Similar results, reported recently in a preprint that has not yet been peer reviewed, also suggest that G614 increases Spike stability and membrane incorporation (Zhang et al., 2020).

TMPRSS2, a type-II transmembrane serine protease, cleaves the viral Spike after receptor binding to enhance entry of MERS-CoV, SARS-CoV, and SARS-CoV-2 (Hoffmann et al., 2020b; Kleine-Weber et al., 2018; Matsuyama et al., 2020; Millet and Whittaker, 2014; Park et al., 2016; Shulla et al., 2011; Zang et al., 2020). Spike 614 is in a pocket adjacent to the fusion peptide near the expected TMPRSS2 cleavage site, suggesting that there could be differences in the propensity and/or requirement for TMPRSS2 of the G614 variant. To test this hypothesis, we infected 293T cells stably expressing the ACE2 receptor in the presence or absence of TMPRSS2 and quantified the titer of infectious virus. We found similar fold changes in titers between D614 and G614 regardless of TMPRSS2 expression (Figure 6A). Hence, entry of G614-bearing viruses into 293T-ACE2 cells compared with D614-bearing viruses is not enhanced by TMPRSS2. Further studies are required to determine whether the G614 variant shows increased titers in lung cells, which may recapitulate native protease expression levels more faithfully, and to determine whether this variant increases the fitness of authentic SARS-CoV-2.

We also tested whether the D614G variations would be similarly neutralized by a polyclonal antibody. Convalescent sera of six San Diego residents, likely infected in early to mid-March 2020, when D614 and G614 were circulating, demonstrate equivalent or better neutralization of a G614-bearing pseudovirus compared with a D614-bearing pseudovirus (Figures 6D and 6E). Although we do not know with which virus each of these individuals were infected, these initial data suggest that, despite increased fitness in cell culture, G614-bearing virions are not intrinsically more resistant to neutralization by convalescent sera.

Additional Sites of Interest in the Spike Gene with Rare Mutations

Spike has very few mutations overall. A small set has reached 0.3% or more of the global population sample, the threshold for automatic tracking at the https://cov.lanl.gov website (Figures 7A and 7B; details are provided in Table S2). Regions in the alignment where entropy is relatively high compared with the rest of Spike (i.e., local clusters of rare mutations) are also tracked (Table S2). Genetic mutations of interest are mapped as amino acid changes onto a Spike structure (Figure 4). The mutation resulting in the signal peptide L5F change recurs many times in the tree and is stably maintained in about 0.6% of the global GISAID data. There are several clusters of mutations in the region of the Spike gene encoding the N-terminal domain (NTD) and RBD that are potential targets for neutralizing antibodies (Chen et al., 2017; Zhou et al., 2019; Sui et al., 2008; Tang et al., 2014; ter Meulen et al., 2006). The RBD cluster (positives 475–483) spans two positions, 475 and 476, that are located within 4 Å of bound ACE2 (Figure 4D; Yan et al., 2020). The fusion peptide contains a cluster of amino acid changes between 826–839; this cluster is highlighted in Figure 7 to illustrate our web-based tools for tracking variation (Figures 7A–7C). The fusion core of HR1 (Xia et al., 2020), next to the helix break in pre-fusion Spike, also contains a cluster of amino acid changes between 936-940 (Figure 4E). The motif SXSS (937–940) may enhance the association of helices (Dawson et al., 2002; Salamango and Johnson, 2015). The cytoplasmic tail of Spike also contains a site of interest, P1263L.

DISCUSSION

Our data show that, over the course of 1 month, the variant carrying the D614G Spike mutation became the globally dominant form of SARS-CoV-2. Phylogenetic tracking of SARS-CoV-2 variants at Nextstrain reveals complex webs of evolutionary and geographical relationships (https://nextstrain.org; Hadfield et al., 2018); travelers globally dispersed G614 variants and likely

Figure 3. Modeling the Daily Fraction of the G614 Variant as a Function of Time in Local Regions Using Isotonic Regression

(A) Analysis summaries for all of the level 3 and 4 regional subdivisions from GISAID data (Figure 1) that have at least 5 each of D614 and G614 variants and that are sampled on at least 14 days. We report the number of each variant, the number of days with test results, and the number of days spanning the first and the last reported tests. the p values are for two one-sided tests, comparing the null hypothesis of no consistent changes in relative frequency over time with positive pressure (the fraction of the G614 variant increasing with time) or negative pressure (the fraction of the G614 variant decreasing with time). Across all regions with sufficient data, binomial p values against the null that increases and decreases are equally likely to indicate that the consistency of increasing G614 is highly significant. California has increasing and decreasing patterns with low p values; this can happen when different time windows support opposing patterns. The G614 decreasing time window in California was driven by sampling from Santa Clara county, a rare region that has retained the D614 form (Figure S4). In the May 29, 2020 dataset used here, Santa Clara county was sampled later in May than any other region in California, so the California G614 frequency dips at this last available time point. When Santa Clara county is removed from the California sample, the pattern of increasing levels of G614 is restored (red asterisk).

(B) Three examples of cities, plotting the daily fraction of G614 as a function of time and accompanied by plots of running weekly counts. The dot size is proportional to the number of sequences sampled that day. The staircase line is the maximum likelihood estimate under the constraint that the logarithm of the odds ratio is non-decreasing. Two typical examples are shown, highlighted in blue (Sydney and Cambridge), and one exception is shown, highlighted in orange (Yakima). Yakima had a brief sampling window enriched for G614 early in the sampling period, but otherwise G614 maintained a low frequency.

Summaries and plots for all regional data at levels 2–4 (included country) are included in Data S1.

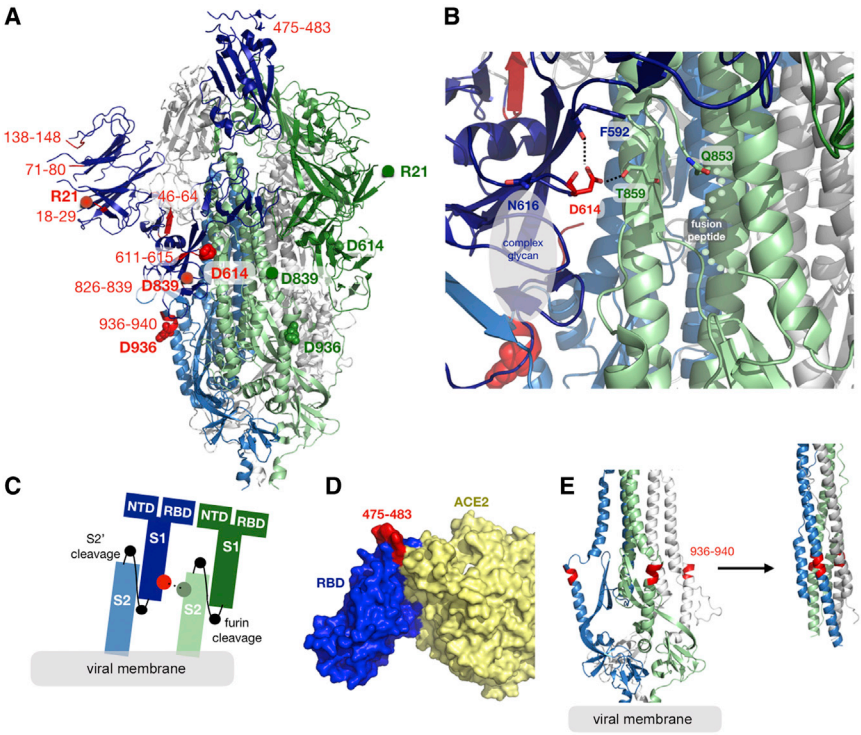


Figure 4. Structural Mapping of Amino Acid Changes and Clusters of Variation in the Spike Protein

(A) Sites including Spike 614 and those noted in Figure 7 mapped onto S_1 and S_2 units of the Spike protein (PDB: 6VSB). S_1 and S_2 are defined based on the furin cleavage site (protomer 1: S_1, dark blue; S_2, light blue; protomer 2: S_1, dark green; S_2, light green; protomer 3: gray). The RBD of protomer 1 is in the "up" position for engagement with the ACE2 receptor. Sites of interest are indicated by red balls, and variable clusters are labeled in red. The missing RBD residues at the ACE2 interface are shown in (D).

(B) Proximity of D614 (red) to an N-linked glycosylation sequon of its own protomer (blue) and to residues T859 and Q853 of the neighboring protomer (green) are shown. Black dashed lines indicate possible hydrogen bond formation. Dotted lines indicate the structurally unresolved region of the fusion peptide connecting to Q853.

(C) Schematic representation of potential protomer-protomer interactions shown in (B), in which D614 (red) from the S_1 unit of one protomer (blue) is brought close to T859 from the S_2 unit of the neighboring protomer.

(D) Sites of interest (red, residues 475–483) near the RBD (blue)-ACE2 (yellow) binding interface. The interfacial region is shown as a molecular surface (PDB: 6M17).

(E) Variable cluster 936–940 (red) in the HR1 region of S_2. These residues occur in a region that undergoes conformational transition during fusion; pre-fusion (PDB: 6VSB) and post-fusion (PDB: 6LXT) conformations of HR1 are shown (left and right).

would have introduced and reintroduced G614 variants into different locations. Still, D614 prevalent epidemics were very well established in many locations when G614 first began to appear (see Figure S2 for examples). The mutation that causes the D614G amino change is transmitted as part of a conserved haplotype defined by 4 mutations that almost always track together (Figures S5 and S6). The pattern of increasing G614 frequency within many different populations where D614 and G614 were co-circulating is highly significant, suggesting that G614 may be under positive selection (Figures 1B and 3). We also found G614 to be associated with higher levels of viral nucleic acid in the upper respiratory tract in human patients (Figure 5), suggestive of higher viral loads, and with higher infectivity in multiple pseudotyping assays (Figure 6).

Given that most G614 variants belong to the G clade lineage, phylogenetic methods that depend on recurrence of mutational events for their signal are poorly powered to resolve whether D614G is under positive selection. The GISAID data, however,

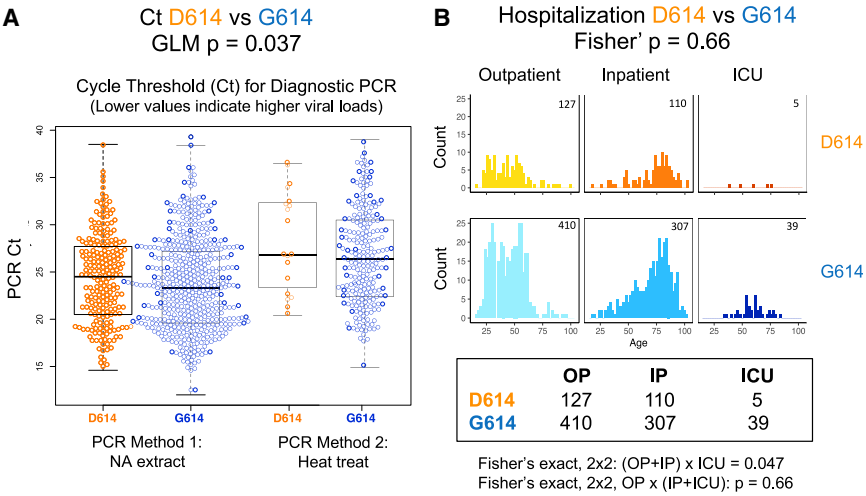


Figure 5. Clinical Status and D614G Associations Based on 999 Subjects with COVID-19 and Linked Sequence and Clinical Data Were Sampled in Sheffield, England

(A) G614 was associated with a lower cycle threshold (Ct) required for detection; lower values are indicative of higher viral loads. The PCR method was changed partway through April 2020 because of shortages of nucleic acid extraction kits (Fomsgaard and Rosenstierne, 2020). The Ct levels for the two PCR methods (nucleic acid extraction versus simple heat inactivation) differ, and so we used a GLM to evaluate the statistical effect of D614G across methods.

(B) D614G status was not statistically associated with hospitalization status (outpatient [OP], inpatient [IP], or ICU) as a marker of disease severity, but age was highly correlated. The number of counts in each category is noted in the top right corner of each graph. See the main text and STAR Methods for statistical details.

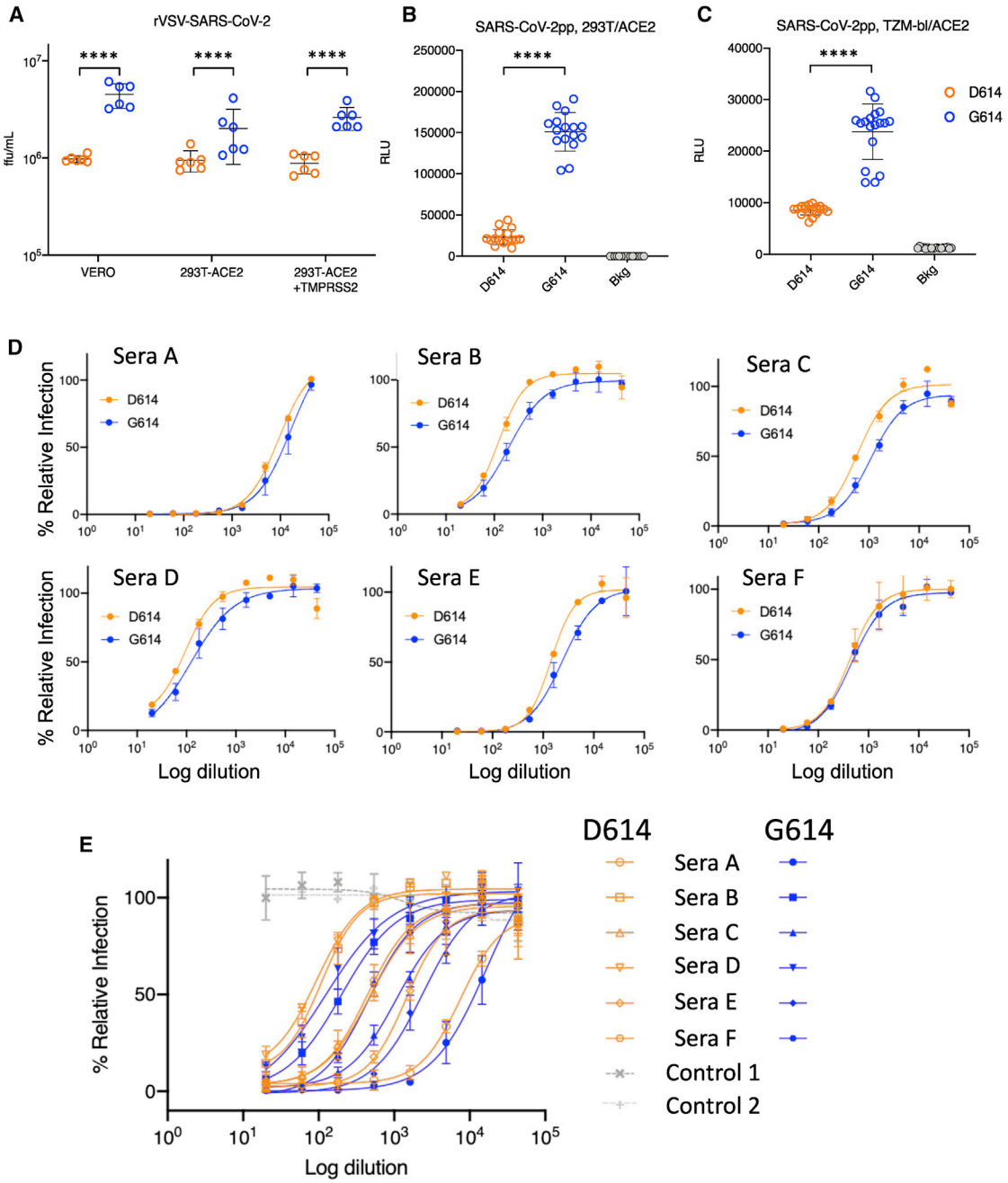


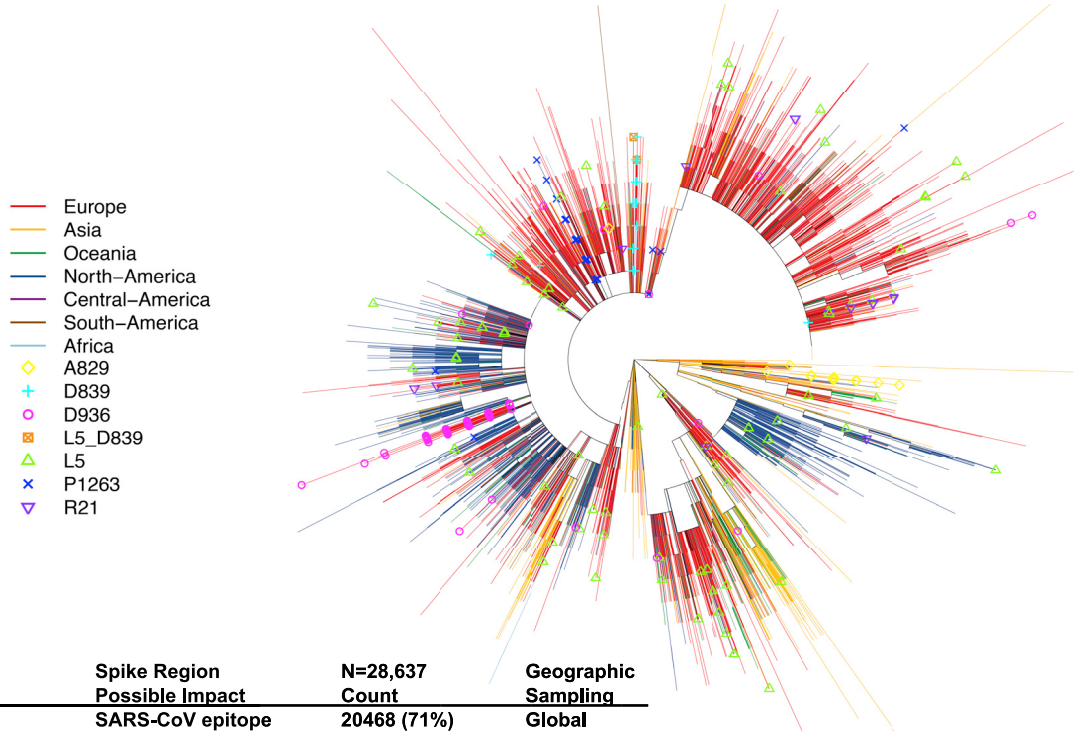
Figure 6. Viral Infectivity and D614G Associations

(A) A recombinant VSV pseudotyped with the G614 Spike grows to a higher titer than D614 Spike in Vero, 293T-ACE2, and 293T-ACE2-TMPRSS2 cells, as measured in terms of focus-forming units (ffu). ****$p < 0.0001$ by Student's t test in pairwise comparisons. Experiments were repeated twice, each time in triplicate. Using a GLM to assess viral infectivity of the D614 and G614 variants across cell types and to account for repeat experiments, we found that the G614 variant had an average 3-fold higher infectious titer than D614 and that this difference was highly significant ($p = 9 \times 10^{-11}$) (STAR Methods).

(B and C) Recombinant lentiviruses pseudotyped with the G614 Spike were more infectious than corresponding D614 S-pseudotyped viruses in (B) 293T/ACE2 (6.5-fold increase) and (C) TZM-bl/ACE2 cells (2.8-fold increase, $p < 0.0001$). Relative luminescence units (RLUs) of Luc reporter gene expression (Naldini et al., 1996) were standardized to the p24 content of the pseudoviruses (p24 content of pseudoviruses for 293T/ACE2 cells: D614 = 269 ng/mL, G614 = 255 ng/mL; p24 content of pseudoviruses for TZM-bl/ACE2 cells, D614 = 680 ng/mL, G614 = 605 ng/mL). Background RLUs were measured in wells that received cells but no pseudovirus.

(D and E) Convalescent serum from six individuals in San Diego (Donors A–F) can neutralize D614-bearing (orange) and G614-bearing (blue) VSV pseudoviruses. Percent relative infection is plotted versus log polyclonal antibody concentration. Error bars indicate the mean ± SD of two biological replicates, each having two technical replicates. In (D), the sensitivity to each form is shown seperately for each sera. In (E) the responses to all sera are combined in one graph, and two negative control normal human sera are indicated in grey.

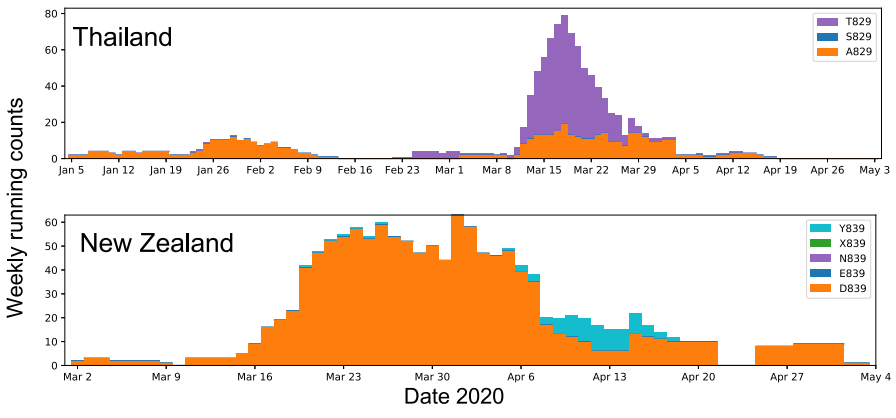
A



B

Spike Mutation	Spike Region Possible Impact	N=28,637 Count	Geographic Sampling
D614G	**SARS-CoV epitope**	**20468 (71%)**	**Global**
L5F	Signal Peptide	170 (0.6%)	Global
R21I/K/T	S1 NTD domain	133 (0.5%)	Wales, England, others
A829T/S	Fusion Peptide	91 (0.3%)	Thailand
D839Y/N/E	Fusion Peptide	149 (0.5%)	Portugal, New Zealand, others
D936Y/H	Heptad Repeat 1(HR1)	257 (0.9%)	Sweden, UK, others
P1263L	Cytoplasmic Tail	201 (0.7%)	UK, others

C



Monitoring sites of interest in local epidemics

D

Higher Entropy Mutational clusters

2-9	Signal Peptide
18-29	N-terminal domain (NTD)
46-54	NTD
71-80	NTD
138-148	NTD
475-483	RBD
611-615	S2, Protomer interactions
826-839	Fusion Peptide
936-940:	HR1

	Count	Percent.	Local cluster
VTLADAGFIKQYGD	28856	98.78	
----------Y	149	0.51	Portugal, New Zealand
---T----------	91	0.31	Thailand
-----V---------	32	0.11	Iceland
L----------	18	0.06	Oxford
----------H---	11	0.04	Cambridge
----------H--	7	0.02	Cambridge
--------T-----	4	0.01	UK
----------L---	3	0.01	Virginia

(legend on next page)

provided the opportunity to look into the relationships among the SARS-CoV-2 variants in the context of time and geography, enabling us to track the increase in frequency of G614 as an early indicator of possible positive selection. This approach is potentially subject to founder effects and sampling biases, and so we generally view this strategy as simply an early indicator of an amino acid change that should be monitored further and tested. The G614 variant stood out, however, in our early detection framework for several reasons. First was the consistency of increase across geographic regions, which was highly significantly non-random (Figures 1B and 3). Second, if the two forms were equally likely to propagate, then one would expect the D614 form to persist in many locations where the G614 form was introduced into the ongoing well-established D614 epidemics. Instead, we found that, even in such cases, G614 increased (Figures 1, 2, 3, S2, and S3). Third, the increase in G614 frequency often continued well after national stay-at-home orders were in place, when serial reseeding from travelers was likely to be reduced significantly (Figures 2, S2, and S3).

Our global tracking data show that the G614 variant in Spike has spread faster than D614. We interpret this to mean that the virus is likely to be more infectious, a hypothesis consistent with the higher infectivity observed with G614 Spike-pseudotyped viruses we observed *in vitro* (Figure 6) and the G614 variant association with higher patient Ct values, indicative of potentially higher *in vivo* viral loads (Figure 5). Interestingly, we did not find evidence of G614 effects on disease severity; i.e., it was not significantly associated with hospitalization status. However, an association between the G614 variant and higher fatality rates has been reported in a comparison of mortality rates across countries, although this kind of analysis can be complicated by different availability of testing and care in different nations (Becerra-Flores and Cardozo, 2020).

Although higher infectiousness of the G614 variant may fully account for its rapid spread and persistence, other factors should also be considered. These include epidemiological factors because viral spread also depends on whom it infects, and epidemiological influences can also cause changes in genotype frequency to mimic evolutionary pressures. In all likelihood, a combination of evolutionary selection for G614 and the founder's effects of being introduced into highly mobile and connected populations may have together contributed, in part, to its rise. The G-clade mutations in the 5′ UTR or in the RdRP protein might also have effects. In addition, there could be immunological consequences resulting from the G614 change in Spike. The G614 variant is sensitive to neutralization by polyclonal convalescent serum (Figure 5), which is encouraging in terms of immune interventions, but it will be important to determine whether the D614 and G614 forms of SARS-CoV-2 are differentially sensitive to neutralization by vaccine-elicited antibodies or by antibodies produced in response to infection with either form of the virus. Also, if the G614 variant is indeed more infectious than the D614 form (Figure 6), then it may require higher antibody levels for protection by vaccines or antibody therapeutic agents than the D614 form. Antibodies against an immunodominant linear epitope spanning Spike 614 in SARS-CoV-1 were associated with ADE activity (Wang et al., 2016), and so it is possible that this mutation may affect ADE.

Tracking mutations in the Spike gene has been our primary focus to date because of its relevance to vaccine and antibody-based therapy strategies currently under development. Such interventions take months to years to develop. For the sake of efficiency, contemporary variation should be factored in during development to ensure that the interventions will be effective against circulating variants when they are eventually deployed. To this end, we built a data analysis pipeline to enable exploration of potentially interesting mutations on SARS-CoV-2 sequences. The analysis is updated daily as the data become available through GISAID, enabling experimentalists to make use of the most current data available to inform vaccine development, reagents for evaluating antibody response, and experimental design. The speed with which the G614 variant became the dominant form globally suggests a need for continued vigilance.

Limitations of Study

Shifts in frequency toward the G614 variant in any given geographic region could, in principle, result from founder effects or sampling biases; it was the consistency of this pattern across regions where both forms of the virus were initially co-circulating that led us to suggest that the G614 form might be transmitted more readily because of an intrinsic fitness advantage; however,

Figure 7. Tracking Variation in Spike

(A) Spike sites of interest (with a minimum frequency of 0.3% variant amino acids) are mapped onto a parsimony tree (for D614G; Figure S6). L5F recurs throughout the tree and is often clustered in small local clades. A829T is found in a single lineage. Other sites of interest cluster in a main lineage but are occasionally found in other parts of the tree in distant geographic regions and, thus, likely to recur at a low level. Build parsimony trees. A brief parsimony search (parsimony ratchet with 5 replicates) is performed with "oblong" (Goloboff, 2014). This is intended as an efficient clustering procedure rather than an explicit attempt to achieve accurate phylogenetic reconstruction, but it appears to yield reasonable results in this situation of a very large number of sequences with a very small number of changes, where more complex models may be subject to overfitting. When multiple most-parsimonious trees are found, only the shortest of these (under a p-distance criterion) is retained. Distance scoring is performed with PAUP* (Swofford, 2003).

(B) The global frequency of amino acid variants in sites of interest and the place where they are most commonly found. Such information could be useful if a vaccine or antibody is intended for use in a geographic region with a commonly circulating variant, so it could be experimentally evaluated prior to testing the planned intervention.

(C) Examples of exploratory plots showing A829T in Thailand and D839Y in New Zealand. Such plots for any variant in any region can be readily created at https://cov.lanl.gov/ to enable monitoring local frequency changes.

(D) Contiguous regions of relatively high entropy in the Spike alignment, indicative of local clusters of amino acid variation in the protein. The fusion-peptide cluster is used as example. It spans two sites of interest, labeled in blue and purple in (B) and (C). The alignment is created with AnalyzeAlign. Contemporary versions of these figures can be created at https://cov.lanl.gov/. Care should be taken to try to avoid systematic sequencing errors and processing artifacts among rare variants (for example, see Figure S7; De Maio et al., 2020; Freeman et al., 2020).

systematic biases across many regions could affect the levels of significance we observed. The lack of association between G614 and hospitalization we report may miss effects on disease severity that are more subtle than we can detect. The experimental approach taken here to acquire laboratory evidence of increased fitness of the D614G mutation is based on two different pseudovirus models of infection in established cell lines. The extent to which this model faithfully recapitulates wild-type virus infection in natural target cells of the respiratory system is still being determined, and our laboratory experiments do not directly address the biology and mechanics of natural transmission. Infectiousness and transmissibility are not always synonymous, and more studies are needed to determine whether the D614G mutation actually led to an increase in number of infections and not just higher viral loads during infection. We encourage others to study this phenomenon in greater detail with a wild-type virus in natural infection and varied target cells (Hou et al., 2020) and in relevant animal models. Finally, the neutralization assays performed were based on sera from SARS-CoV-2 infected individuals with an unknown D614G status. Thus, although they show that the G614 variants are neutralization sensitive, more work is needed to resolve whether the potency of neutralization is affected when the variant that initiated the immune response differs from the test variant or when monoclonal antibodies are used.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human Subjects
- METHOD DETAILS
  - Detection and Sequencing of Sars-Cov-2 Isolates from Clinical Samples
  - Pseudotyped Virus Infectivity
  - Data Pipeline
  - Global Maps
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Systematic Regional Analysis of D614/G614 Frequencies
  - Clinical Data and Modeling
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2020.06.043.

CONSORTIA

The members of the Sheffield COVID-19 Genomics Group are Adrienne Angyal, Rebecca L. Brown, Laura Carrilero, Katie Cooke, Alison Cope, Thushan I. de Silva, Mark Dunning, Cariad M. Evans, Luke R. Green, Danielle C. Groves, Hailey Hornsby, Katie J. Johnson, Sokratis Karitois, Alexander J. Keeley, Katjusa Koler, Benjamin B. Lindsey, Matthew D. Parker, Paul J. Parsons, David G. Partridge, Mohammad Raza, Sarah Rowland-Jones, Nikki Smith, Rachel M. Tucker, Dennis Wang, and Matthew D. Wyles.

ACKNOWLEDGMENTS

We thank Andrew McMichael, Sarah Rowland-Jones, and Xiao-Ning Xu for bringing together the clinical and theory teams. We thank Anthony West for pointing out the 5′ UTR G-clade mutation; George Ellison for suggestions regarding clinical data analyses; Barbara Imperiali for insights regarding the structural implications of the D614G change; and Rachael Mansbach, Srirupa Chakraborty, and Kien Nguyen for sharing preliminary MD data. We thank Davey Smith and Stephen Rawlings of UCSD and Alessandro Sette, Jennifer M. Dan, and Shane Crotty of LJI for survivor sera and Sharon Schendel for manuscript edits. We acknowledge Barney Graham, Kizzmekia Corbett, Nicole Doria-Rose, Adrian McDermott, and John Mascola at the Vaccine Research Center, NIH for reagents and assistance with the lentivirus-based SARS-CoV-2 infection assay and Elize Domin for technical support. The Sheffield COVID-19 Genomics Group is part of the COG-UK CONSORTIUM, supported by the Medical Research Council (MRC), part of UK Research & Innovation (UKRI), the National Institute of Health Research (NIHR), and Genome Research Limited, operating as the Wellcome Sanger Institute. T.I.d.S. is supported by a Wellcome Trust intermediate clinical fellowship (110058/Z/15/Z). M.P. was funded by the NIHR Sheffield Biomedical Research Centre (BRC). B.K., E.E.G., T.B., N.H., W.M.F., H.Y., and W.A. were supported by LANL LDRD projects 20200554ECR and 20200706ER and through NIH NIAID, DHHS Interagency Agreement AAI12007-001-00000. D.S. acknowledges support from the John and Mary Tu Foundation and the San Diego CFAR AI036214. A.S. and S.C. acknowledge support from NIH NIAID AI42742 (Cooperative Centers for Human Immunology). E.O.S. acknowledges support from CoVIC INV-006133 of the COVID-19 Therapeutics Accelerator, supported by the Bill and Melinda Gates Foundation, Mastercard, Wellcome and private philanthropic support, the Overton family, and a FastGrant from Emergent Ventures in aid of COVID-19 research. We gratefully acknowledge the team at GISAID for creating SARS-CoV-2 global database and the many people who provided sequence data (Table S1).

AUTHOR CONTRIBUTIONS

Conceptualization, B.K. and D.C.M.; Methodology, B.K., W.M.F., J.T., N.H., E.O.S., and D.C.M.; Software, W.M.F., J.T., H.Y., W.A., N.H., E.E.G., T.B., T.M.F., M.D.P., and B.K.; Validation, E.O.S., D.C.M., J.T., B.K., B.F., and N.H.; Formal Analysis, B.K., J.T., N.H., W.M.F., S.G., M.D.P., T.M.F., D.G.P., C.M.E., T.I.d.S., T.B., and E.E.G.; Investigation, E.O.S., D.C.M., K.M.H., C.M.E., D.G.P., L.G.P., H.T., A.M.-W., S.P.W., C.C.L., and T.I.d.S.; Writing – Original Draft, B.K., W.M.F., S.G., D.C.M., and E.O.S; Writing – Review & Editing, T.I.d.S., C.C.L., E.E.G., N.H., H.Y., and T.B.; Visualization, B.K., E.O.S., J.T., N.H., W.M.F., E.E.G., and S.G.; Supervision, B.K., D.C.M., E.O.S., T.I.d.S., and S.P.W.; Funding Acquisition, B.K., E.E.G., D.C.M., S.G., E.O.S., and T.I.d.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 29, 2020
Revised: June 10, 2020
Accepted: June 26, 2020
Published: July 3, 2020

REFERENCES

Asmal, M., Hellmann, I., Liu, W., Keele, B.F., Perelson, A.S., Bhattacharya, T., Gnanakaran, S., Daniels, M., Haynes, B.F., Korber, B.T., et al. (2011). A signature in HIV-1 envelope leader peptide associated with transition from acute to

chronic infection impacts envelope processing and infectivity. PLoS ONE *6*, e23673.

Becerra-Flores, M., and Cardozo, T. (2020). SARS-CoV-2 viral spike G614 mutation exhibits higher case fatality rate. Int. J. Clin. Pract. Published online May 6, 2020. https://doi.org/10.1111/ijcp.13525.

Boni, M.F., Gog, J.R., Andreasen, V., and Feldman, M.W. (2006). Epidemic dynamics and antigenic evolution in a single season of influenza A. Proc. Biol. Sci. *273*, 1307–1316.

Bricault, C.A., Yusim, K., Seaman, M.S., Yoon, H., Theiler, J., Giorgi, E.E., Wagh, K., Theiler, M., Hraber, P., Macke, J.P., et al. (2019). HIV-1 Neutralizing Antibody Signatures and Application to Epitope-Targeted Vaccine Design. Cell Host Microbe *26*, 296.

Chen, W.H., Hotez, P.J., and Bottazzi, M.E. (2020). Potential for developing a SARS-CoV receptor-binding domain (RBD) recombinant protein as a heterologous human vaccine against coronavirus infectious disease (COVID)-19. Hum. Vaccin. Immunother. *16*, 1239–1242.

Chen, Y., Lu, S., Jia, H., Deng, Y., Zhou, J., Huang, B., Yu, Y., Lan, J., Wang, W., Lou, Y., et al. (2017). A novel neutralizing monoclonal antibody targeting the N-terminal domain of the MERS-CoV spike protein. Emerging Microbes and Infections *6*, e37.

Chibo, D., and Birch, C. (2006). Analysis of human coronavirus 229E spike and nucleoprotein genes demonstrates genetic drift between chronologically distinct strains. J. Gen. Virol. *87*, 1203–1208.

Cohen, J. (2020). COVID-19 shot protects monkeys. Science *368*, 456–457.

Conti, P., and Younes, A. (2020). Coronavirus COV-19/SARS-CoV-2 affects women less than men: clinical response to viral infection. J. Biol. Regul. Homeost. Agents *34*. https://doi.org/10.23812/Editorial-Conti-3.

Corman, V.M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D.K., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M.L., et al. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. *25*, 2000045.

Crispell, J., Balaz, D., and Gordon, S.V. (2019). HomoplasyFinder: a simple tool to identify homoplasies on a phylogeny. Microb. Genom. *5*, e000245.

Cui, J., Li, F., and Shi, Z.-L. (2019). Origin and evolution of pathogenic coronaviruses. Nat. Rev. Microbiol. *17*, 181–192.

Dawson, J.P., Weinger, J.S., and Engelman, D.M. (2002). Motifs of serine and threonine can drive association of transmembrane helices. J. Mol. Biol. *316*, 799–805.

De Maio, N., Walker, C., Borges, R., Weilguny, L., Slodkowicz, G., and Goldman, N. (2020). Issues with SARS-CoV-2 sequencing data. https://virological.org/t/issues-with-sars-cov-2-sequencing-data/473.

de Wit, E., van Doremalen, N., Falzarano, D., and Munster, V.J. (2016). SARS and MERS: recent insights into emerging coronaviruses. Nat. Rev. Microbiol. *14*, 523–534.

Deng, X., Gu, W., Federman, S., du Plessis, L., Pybus, O.G., Faria, N., Wang, C., Yu, G., Bushnell, B., Pan, C.Y., et al. (2020). Genomic surveillance reveals multiple introductions of SARS-CoV-2 into Northern California. Science, eabb9263.

Dowd, J.B., Andriano, L., Brazel, D.M., Rotondi, V., Block, P., Ding, X., Liu, Y., and Mills, M.C. (2020). Demographic science aids in understanding the spread and fatality rates of COVID-19. Proc. Natl. Acad. Sci. USA *117*, 9696–9698.

Elbe, S., and Buckland-Merrett, G. (2017). Data, disease and diplomacy: GISAID's innovative contribution to global health. Glob. Chall. *1*, 33–46.

Fauver, J.R., Petrone, M.E., Hodcroft, E.B., Shioda, K., Ehrlich, H.Y., Watts, A.G., Vogels, C.B.F., Brito, A.F., Alpert, T., Muyombwe, A., et al. (2020). Coast-to-Coast Spread of SARS-CoV-2 during the Early Epidemic in the United States. Cell *181*, 990–996.e5.

Fomsgaard, A.S., and Rosenstierne, M.W. (2020). An alternative workflow for molecular detection of SARS-CoV-2 - escape from the NA extraction kit-shortage, Copenhagen, Denmark, March 2020. Euro Surveill. *25*, 2000398.

Freeman, T.M., Wang, D., and Harris, J.; Genomics England Research Consortium (2020). Genomic loci susceptible to systematic sequencing bias in clinical whole genomes. Genome Res. *30*, 415–426.

Gao, F., Bonsignori, M., Liao, H.X., Kumar, A., Xia, S.M., Lu, X., Cai, F., Hwang, K.K., Song, H., Zhou, T., et al. (2014). Cooperation of B cell lineages in induction of HIV-1-broadly neutralizing antibodies. Cell *158*, 481–491.

Goloboff, P.A. (2014). Oblong, a program to analyse phylogenomic data sets with millions of characters, requiring negligible amounts of RAM. Cladistics *30*, 273–281.

Gorbalenya, A.E., Baker, S.C., Baric, R.S., de Groot, R.J., Drosten, C., Gulyaeva, A.A., Haagmans, B.L., Lauber, C., Leontovich, A.M., Neuman, B.W., et al.; Coronaviridae Study Group of the International Committee on Taxonomy of Viruses (2020). The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat. Microbiol. *5*, 536–544.

Graham, R.L., and Baric, R.S. (2010). Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. J. Virol. *84*, 3134–3146.

Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L., Luo, S.W., Li, P.H., Zhang, L.J., Guan, Y.J., et al. (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science *302*, 276–278.

Gudbjartsson, D.F., Helgason, A., Jonsson, H., Magnusson, O.T., Melsted, P., Norddahl, G.L., Saemundsdottir, J., Sigurdsson, A., Sulem, P., Agustsdottir, A.B., et al. (2020). Spread of SARS-CoV-2 in the Icelandic Population. N. Engl. J. Med. *382*, 2302–2315.

Hadfield, J., Megill, C., Bell, S.M., Huddleston, J., Potter, B., Callender, C., Sagulenko, P., Bedford, T., and Neher, R.A. (2018). Nextstrain: real-time tracking of pathogen evolution. Bioinformatics *34*, 4121–4123.

Ho, M.-S., Chen, W.-J., Chen, H.-Y., Lin, S.-F., Wang, M.-C., Di, J., Lu, Y.-T., Liu, C.-L., Chang, S.-C., Chao, C.-L., et al. (2005). Neutralizing antibody response and SARS severity. Emerg. Infect. Dis. *11*, 1730–1737.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.-H., Nitsche, A., et al. (2020a). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell *181*, 271–280.e8.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A., et al. (2020b). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell *181*, 271–280.e8.

Hou, Y.J., Okuda, K., Edwards, C.E., Martinez, D.R., Asakura, T., Dinnon, K.H., III, Kato, T., Lee, R.E., Yount, B.L., Mascenik, T.M., et al. (2020). SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. Cell. Published online May 27, 2020. https://doi.org/10.1016/j.cell.2020.05.042.

Hunter, J.D. (2007). Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. *9*, 90–95.

Jaume, M., Yip, M.S., Cheung, C.Y., Leung, H.L., Li, P.H., Kien, F., Dutry, I., Callendret, B., Escriou, N., Altmeyer, R., et al. (2011). Anti-severe acute respiratory syndrome coronavirus spike antibodies trigger infection of human immune cells via a pH- and cysteine protease-independent FcγR pathway. J. Virol. *85*, 10582–10597.

Kleine-Weber, H., Elzayat, M.T., Hoffmann, M., and Pöhlmann, S. (2018). Functional analysis of potential cleavage sites in the MERS-coronavirus spike protein. Sci. Rep. *8*, 16597.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004). Versatile and open software for comparing large genomes. Genome Biol. *5*, R12.

LaBranche, C.C., Henderson, R., Hsu, A., Behrens, S., Chen, X., Zhou, T., Wiehe, K., Saunders, K.O., Alam, S.M., Bonsignori, M., et al. (2019). Neutralization-guided design of HIV-1 envelope trimers with high affinity for the unmutated common ancestor of CH235 lineage CD4bs broadly neutralizing antibodies. PLoS Pathog. *15*, e1008026.

Lau, S.K., Lee, P., Tsang, A.K., Yip, C.C., Tse, H., Lee, R.A., So, L.Y., Lau, Y.L., Chan, K.H., Woo, P.C., et al. (2011). Molecular epidemiology of human coronavirus OC43 reveals evolution of different genotypes over time and recent emergence of a novel genotype due to natural recombination. J. Virol. *85*, 11325–11337.

Li, H., Wang, S., Kong, R., Ding, W., Lee, F.H., Parker, Z., Kim, E., Learn, G.H., Hahn, P., Policicchio, B., et al. (2016). Envelope residue 375 substitutions in simian-human immunodeficiency viruses enhance CD4 binding and replication in rhesus macaques. Proc. Natl. Acad. Sci. USA *113*, E3413–E3422.

Li, X., Giorgi, E.E., Marichannegowda, M.H., Foley, B., Xiao, C., Kong, X.-P., Chen, Y., Gnanakaran, S., Korber, B., and Gao, F. (2020). Emergence of SARS-CoV-2 through recombination and strong purifying selection. Science Advances *6*, eabb9153.

Liu, W., Fontanet, A., Zhang, P.-H., Zhan, L., Xin, Z.-T., Baril, L., Tang, F., Lv, H., and Cao, W.-C. (2006). Two-year prospective study of the humoral immune response of patients with severe acute respiratory syndrome. J. Infect. Dis. *193*, 792–795.

Liu, Y., Yang, Y., Zhang, C., Huang, F., Wang, F., Yuan, J., Wang, Z., Li, J., Li, J., Feng, C., et al. (2020). Clinical and biochemical indexes from 2019-nCoV infected patients linked to viral loads and lung injury. Sci. China Life Sci. *63*, 364–374.

Lorenzo-Redondo, R., Nam, H.H., Roberts, S.C., Simons, L.M., Jennings, L.J., Qi, C., Achenbach, C.J., Hauser, A.R., Ison, M.G., Hultquist, J.F., et al. (2020). A Unique Clade of SARS-CoV-2 Viruses is Associated with Lower Viral Loads in Patient Upper Airways. medRxiv. https://doi.org/10.1101/2020.2005.2019.20107144.

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., et al. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet *395*, 565–574.

Lv, M., Luo, X., Estill, J., Liu, Y., Ren, M., Wang, J., Wang, Q., Zhao, S., Wang, X., Yang, S., et al.; On Behalf Of The Covid-Evidence And Recommendations Working Group (2020). Coronavirus disease (COVID-19): a scoping review. Euro Surveill. *25*, 2000125.

Matsuyama, S., Nao, N., Shirato, K., Kawase, M., Saito, S., Takayama, I., Nagata, N., Sekizuka, T., Katoh, H., Kato, F., et al. (2020). Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. Proc. Natl. Acad. Sci. USA *117*, 7001–7003.

Millet, J.K., and Whittaker, G.R. (2014). Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. Proc. Natl. Acad. Sci. USA *111*, 15214–15219.

Naldini, L., Blömer, U., Gage, F.H., Trono, D., and Verma, I.M. (1996). Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. Proc. Natl. Acad. Sci. USA *93*, 11382–11388.

Oong, X.Y., Ng, K.T., Takebe, Y., Ng, L.J., Chan, K.G., Chook, J.B., Kamarulzaman, A., and Tee, K.K. (2017). Identification and evolutionary dynamics of two novel human coronavirus OC43 genotypes associated with acute respiratory infections: phylogenetic, spatiotemporal and transmission network analyses. Emerg. Microbes Infect. *6*, e3.

Park, J.E., Li, K., Barlan, A., Fehr, A.R., Perlman, S., McCray, P.B., Jr., and Gallagher, T. (2016). Proteolytic processing of Middle East respiratory syndrome coronavirus spikes expands virus tropism. Proc. Natl. Acad. Sci. USA *113*, 12262–12267.

Promislow, D.E.L. (2020). A geroscience perspective on COVID-19 mortality. J. Gerontol. A Biol. Sci. Med. Sci. Published online April 17, 2020. https://doi.org/10.1093/gerona/glaa094.

Rehman, S., Sharique, L., Ihsan, A., and Liu, Q. (2020). Evolutionary Trajectory for the Emergence of Novel Coronavirus SARS-CoV-2. Pathogens *9*, 240.

Ren, L., Zhang, Y., Li, J., Xiao, Y., Zhang, J., Wang, Y., Chen, L., Paranhos-Baccalà, G., and Wang, J. (2015). Genetic drift of human coronavirus OC43 spike gene during adaptive evolution. Sci. Rep. *5*, 11451.

Sadjadpour, R., Donau, O.K., Shingai, M., Buckler-White, A., Kao, S., Strebel, K., Nishimura, Y., and Martin, M.A. (2013). Emergence of gp120 V3 variants confers neutralization resistance in an R5 simian-human immunodeficiency virus-infected macaque elite neutralizer that targets the N332 glycan of the human immunodeficiency virus type 1 envelope glycoprotein. J. Virol. *87*, 8798–8804.

Salamango, D.J., and Johnson, M.C. (2015). Characterizing the Murine Leukemia Virus Envelope Glycoprotein Membrane-Spanning Domain for Its Roles in Interface Alignment and Fusogenicity. J. Virol. *89*, 12492–12500.

Sehra, S.T., Salciccioli, J.D., Wiebe, D.J., Fundin, S., and Baker, J.F. (2020). Maximum Daily Temperature, Precipitation, Ultra-Violet Light and Rates of Transmission of SARS-Cov-2 in the United States. Clin. Infect. Dis., ciaa681.

Sevajol, M., Subissi, L., Decroly, E., Canard, B., and Imbert, I. (2014). Insights into RNA synthesis, capping, and proofreading mechanisms of SARS-coronavirus. Virus Res. *194*, 90–99.

Shu, Y., and McCauley, J. (2017). GISAID: Global initiative on sharing all influenza data - from vision to reality. Euro Surveill. *22*, 30494.

Shulla, A., Heald-Sargent, T., Subramanya, G., Zhao, J., Perlman, S., and Gallagher, T. (2011). A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. J. Virol. *85*, 873–882.

Smith, E.C., Blanc, H., Surdel, M.C., Vignuzzi, M., and Denison, M.R. (2013). Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. PLoS Pathog. *9*, e1003565.

Song, H.D., Tu, C.C., Zhang, G.W., Wang, S.Y., Zheng, K., Lei, L.C., Chen, Q.X., Gao, Y.W., Zhou, H.Q., Xiang, H., et al. (2005). Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. Proc. Natl. Acad. Sci. USA *102*, 2430–2435.

Sui, J., Aird, D.R., Tamin, A., Murakami, A., Yan, M., Yammanuru, A., Jing, H., Kan, B., Liu, X., Zhu, Q., et al. (2008). Broadening of neutralization activity to directly block a dominant antibody-driven SARS-coronavirus evolution pathway. PLoS Pathog. *4*, e1000197.

Swofford, D.L. (2003). PAUP* {Phylogenetic Analysis Using Parsimony (and Other Methods)], Version 4. In PAUP* Phylogenetic Analysis Using Parsimony (Sinauer Associates).

Tang, X.-C., Agnihothram, S.S., Jiao, Y., Stanhope, J., Graham, R.L., Peterson, E.C., Avnir, Y., Tallarico, A.S.C., Sheehan, J., Zhu, Q., et al. (2014). Identification of human neutralizing antibodies against MERS-CoV and their role in virus adaptive evolution. Proc. Natl. Acad. Sci. USA *111*, E2018–E2026.

Temperton, N.J., Chan, P.K., Simmons, G., Zambon, M.C., Tedder, R.S., Takeuchi, Y., and Weiss, R.A. (2005). Longitudinally profiling neutralizing antibody response to SARS coronavirus with pseudotypes. Emerg. Infect. Dis. *11*, 411–416.

ter Meulen, J., van den Brink, E.N., Poon, L.L., Marissen, W.E., Leung, C.S., Cox, F., Cheung, C.Y., Bakker, A.Q., Bogaards, J.A., van Deventer, E., et al. (2006). Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants. PLoS Med. *3*, e237.

Vijgen, L., Keyaerts, E., Lemey, P., Moës, E., Li, S., Vandamme, A.-M., and Van Ranst, M. (2005). Circulation of genetically distinct contemporary human coronavirus OC43 strains. Virology *337*, 85–92.

Wagner, C., Roychoudhury, P., Hadfield, J., Hodcroft, E., Lee, J., Moncla, L., Muller, N., Behrens, C., Huang, M.-L., Mathias, P., et al. (2020). Comparing viral load and clinical outcomes in Washington State across D614G mutation in spike protein of SARS-CoV-2. https://github.com/blab/ncov-D614G.

Walls, A.C., Park, Y.-J., Tortorici, M.A., Wall, A., McGuire, A.T., and Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell *181*, 281–292.e6.

Wan, Y., Shang, J., Sun, S., Tai, W., Chen, J., Geng, Q., He, L., Chen, Y., Wu, J., Shi, Z., et al. (2020). Molecular Mechanism for Antibody-Dependent Enhancement of Coronavirus Entry. J. Virol. *94*, e02015–e02019.

Wang, S.-F., Tseng, S.-P., Yen, C.-H., Yang, J.-Y., Tsao, C.-H., Shen, C.-W., Chen, K.-H., Liu, F.-T., Liu, W.-T., Chen, Y.-M.A., and Huang, J.C. (2014).

Antibody-dependent SARS coronavirus infection is mediated by antibodies against spike proteins. Biochem. Biophys. Res. Commun. *451*, 208–214.

Wang, Q., Zhang, L., Kuwahara, K., Li, L., Liu, Z., Li, T., Zhu, H., Liu, J., Xu, Y., Xie, J., et al. (2016). Immunodominant SARS Coronavirus Epitopes in Humans Elicited both Enhancing and Neutralizing Effects on Infection in Non-human Primates. ACS Infect. Dis. *2*, 361–376.

Wang, C., Liu, Z., Chen, Z., Huang, X., Xu, M., He, T., and Zhang, Z. (2020). The establishment of reference sequence for SARS-CoV-2 and variation analysis. J. Med. Virol. *92*, 667–674.

Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.-L., Abiona, O., Graham, B.S., and McLellan, J.S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science *367*, 1260–1263.

Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., Meng, J., Zhu, Z., Zhang, Z., Wang, J., et al. (2020). Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. Cell Host Microbe *27*, 325–328.

Xia, S., Liu, M., Wang, C., Xu, W., Lan, Q., Feng, S., Qi, F., Bao, L., Du, L., Liu, S., et al. (2020). Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. Cell Res. *30*, 343–355.

Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., and Zhou, Q. (2020). Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science *367*, 1444–1448.

Yip, M.S., Leung, H.L., Li, P.H., Cheung, C.Y., Dutry, I., Li, D., Daëron, M., Bruzzone, R., Peiris, J.S., and Jaume, M. (2016). Antibody-dependent enhancement of SARS coronavirus infection and its role in the pathogenesis of SARS. Hong Kong Med. J. *22* (3, Suppl 4), 25–31.

Yu, J., Tostanoski, L.H., Peter, L., Mercado, N.B., McMahan, K., Mahrokhian, S.H., Nkolola, J.P., Liu, J., Li, Z., Chandrashekar, A., et al. (2020). DNA vaccine protection against SARS-CoV-2 in rhesus macaques. Science, eabc6284.

Yuan, M., Wu, N.C., Zhu, X., Lee, C.-C.D., So, R.T.Y., Lv, H., Mok, C.K.P., and Wilson, I.A. (2020). A highly conserved cryptic epitope in the receptor-binding domains of SARS-CoV-2 and SARS-CoV. Science, eabb7269.

Zang, R., Gomez Castro, M.F., McCune, B.T., Zeng, Q., Rothlauf, P.W., Sonnek, N.M., Liu, Z., Brulois, K.F., Wang, X., Greenberg, H.B., et al. (2020). TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. Sci. Immunol. *5*, eabc3582.

Zhang, L., Zhang, F., Yu, W., He, T., Yu, J., Yi, C.E., Ba, L., Li, W., Farzan, M., Chen, Z., et al. (2006). Antibody responses against SARS coronavirus are correlated with disease outcome of infected individuals. J. Med. Virol. *78*, 1–8.

Zhang, L., Jackson, C.B., Mou, H., Ojha, A., Rangarajan, E.S., Izard, T., Farzan, M., and Choe, H. (2020). The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. https://www.scripps.edu/news-and-events/press-room/2020/20200611-choe-farzan-sars-cov-2-spike-protein.html.

Zhou, P., Wang, H., Fang, M., Li, Y., Wang, H., Shi, S., Li, Z., Wu, J., Han, X., Shi, X., et al. (2019). Broadly resistant HIV-1 against CD4-binding site neutralizing antibodies. PLoS Pathog. *15*, e1007819.

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature *579*, 270–273.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Polyclonal human sera	This study	N/A
Bacterial and Virus Strains		
VSV-ΔG-GFP	Karafast	Cat# EH1020
rVSV-SARS-CoV-2	This study	N/A
rVSV-SARS-CoV-2 – D614G	This study	N/A
Chemicals, Peptides, and Recombinant Proteins		
Fugene 6	Promega	Cat# E2692
TransIT-LT1	Mirus	Cat# MIR 2304
PFA	Electron Microscopy Sciences	Cat# 15710
Hoechst	ThermoFisher Scientific	Cat# 62249
SuperScript IV (50rxn	ThermoFisher Scientific	Cat# 18090050
dNTP mix (10mM each)	ThermoFisher Scientific	Cat# R0192
Random hexamers (50uM)	ThermoFisher Scientific	Cat# N8080127
RNase OUT	ThermoFisher Scientific	Cat# 10777019
Q5 High-fidelity polymerase	New England Biolabs	Cat# M0491S
Critical Commercial Assays		
Promega Luciferase Assay System	Promega	Cat# E1501
Britelite Plus Reporter Gene Assay System	Perkin-Elmer Part	Cat# 6066769
MagnaPure96 extraction platform	Roche Diagnostics Ltd, Burgess Hill, UK	Cat# 06 543 588 001
SensiFASTTM Probe No-ROX One-Step Real-time PCR kit	Bioline	Cat# BIO-76001
Ligation sequencing kit	Oxford Nanopore	Cat# SQK-LSK109
Native barcoding expansion kit 13-24	Oxford Nanopore	Cat# EXP-NBD114
Native barcoding expansion kit1-12	Oxford Nanopore	Cat# EXP-NBD104
Flow cell priming kit	Oxford Nanopore	Cat# EXP-FLP002
Flow cells R9.4.1 48pk	Oxford Nanopore	Cat# FLO-MIN106D
Flow cell wash kit	Oxford Nanopore	Cat# EXP-WSH003
SFP Expansion kit	Oxford Nanopore	Cat# EXP-SFB001
Next Ultra II library prep kit (illumina)	New England Biolabs	Cat# E7645L
Quick ligation module	New England Biolabs	Cat# E6056L
HIV-1 p24 ELISA	Perkin-Elmer	Cat# NEK050B
Deposited Data		
COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE)	Johns Hopkins University	https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_US.csv
GISAID	Elbe and Buckland-Merrett, 2017; Shu and McCauley, 2017	https://www.gisaid.org
STAY-AT-HOME ORDERS IN EUROPE		https://www.sidley.com/-/media/uploads/stay-at-home-tracker_europe.pdf?la=en
Experimental Models: Cell Lines		
HEK293T/17 cells	ATCC	Cat# CRL-11268
293T/ACE2.MF	Dr. Michael Farzan and Huihui Mu	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
TZM-bl/ACE2.MF	Dr. Michael Farzan and Huihui Mu	N/A
293T	ATCC	Cat# CRL-3216
Vero	ATCC	Cat# CCL-81
293T-Ace2	This study	N/A
293T-Ace2-TMPRSS2	This study	N/A
Oligonucleotides		
Primer pool1 and Primer Pool2 for sequencing (98 oligonucleotides	Artic Network	https://artic.network/ncov-2019
Recombinant DNA		
VRC7480 (Spike plasmid)	Drs. Barney Graham and Kizzmekia Corbett	N/A
VRC7480.D614G (Spike plasmid)	This study	N/A
pCMV ΔR8.2 (lentiviral backbone)	Drs. Barney Graham and Kizzmekia Corbett	N/A
pHR' CMV Luc (luciferase reporter)	Drs. Barney Graham and Kizzmekia Corbett; Naldini et al., 1996	N/A
pSG3ΔEnv	Drs. Beatrice Hahn and Feng Gao	N/A
Empty vector: phCMV3	Genlantis	Cat# P003300
pCAGGS-VSV-G	Kerfast	Cat# EH1017
phCMV3-SARS-CoV-2	This study, Spike cloned from synthetic, codon optimized DNA	N/A
phCMV3-SARS-CoV-2 – D614G	This study, generated through site-directed mutagenesis	N/A
Software and Algorithms		
R	The R Foundation for Statistical Computing	http://www.R-project.org
Nanopolish	© Ontario Institute for Cancer Research 2015 MPL liscense	https://github.com/jts/nanopolish
R packages: phangorn (version 2.5.5), ggplot2 (version 3.3.0), beeswarm (version 0.2.0), tidyverse (version 1.3.0), ape (version 5.3), lme4 (version 1.1.21)	The R Foundation for Statistical Computing	https://cran.r-project.org/
data.table (version 1.12.8)	Matt Dowle	https://github.com/Rdatatable/data.table
Aliview	Anders Larsson	https://ormbunkar.se/aliview/
cgam (version 1.14)	Xiyue Liao, Mary C. Meyer	https://www.jstatsoft.org/htaccess.php?volume=089&type=i&issue=05
ARTIC network protocol (accessed the 19th of April)	ARTIC network	https://artic.network/ncov-2019
Python Matplotlib A 2D Graphics Environment v 3.2.2	Hunter, 2007	https://matplotlib.org
The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC	Schrödinger	https://pymol.org/2/
Oblong	Goloboff, 2014	http://www.lillo.org.ar/phylogeny/oblong/
Los Alamos National Lab HIV Database: Analyze Align and Entropy	Los Alamos National Lab	http://cov.lanl.govcontent/index
Los Alamos National Lab SARS-CoV-2 Analysis Pipeline: SARS-CoV-2 map, Relative Frequency Change by Geographical Region, Rainbow Tree	This study	http://cov.lanl.govcontent/index
PAUP	David Swofford	https://paup.phylosolutions.com
GraphPad Prism 8	GraphPad Software, Inc	https://www.graphpad.com

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Other		
Published primers and probes for the SARS-CoV-2 E-gene RT- qPCR SARS-CoV-2	Corman et al., 2020	N/A
ABI Thermal Cycler	Applied Biosystems, Foster City, United States	Cat# 4375305
CellInsight CX5 High Content Screening Platform	ThermoFisher Scientific	Cat# CX51110

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Bette Korber (btk@lanl.gov).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

All sequence data used here are available from The Global Initiative for Sharing All Influenza Data (GISAID), at https://www.gisaid.org/. The user agreement for GISAID does not permit redistribution of sequences. Other data have been deposited to Mendeley Data: https://doi.org/10.17632/hn3h9gdrgj.1.

Web-based tools to recreate much of the analyses provided in this paper but based on contemporary GISAID data downloads are available at https://cov.lanl.gov/.

Code to create the alignments as described in Figure S1 and to perform the Isotonic regression analysis in Figure 3 will be available through https://cov.lanl.gov, at also GitHub, once permission from our funders is obtained.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human Subjects

999 individuals presenting with active COVID-19 disease were sampled for SARS CoV-2 sequencing at Sheffield Teaching Hospitals NHS Foundation Trust, UK using samples collected for routine clinical diagnostic use. This work was performed under approval by the Public Health England Research Ethics and Governance Group for the COVID-19 Genomics UK consortium (R&D NR0195). SARS-CoV-2 sequences were generated using samples taken for routine clinical diagnostic use from 999 individuals presenting with active COVID-19 disease: 593 female, 399 male, 6 no gender specified; ages 15-103 (median 55) years.

METHOD DETAILS

Detection and Sequencing of Sars-Cov-2 Isolates from Clinical Samples

Samples for PCR detection of SARS-CoV-2 (Figure 5A) were all obtained from either throat or combined nose/throat swabs. Nucleic acid was extracted from 200μl of sample on MagnaPure96 extraction platform (Roche Diagnostics Ltd, Burgess Hill, UK). SARS-CoV-2 RNA was detected using primers and probes targeting the E gene and the RdRp genes for routine clinical diagnostic purposes, with thermocycling and fluorescence detection on ABI Thermal Cycler (Applied Biosystems, Foster City, United States) using previously described primer and probe sets (Corman et al., 2020). Nucleic acid from positive cases underwent long-read whole genome sequencing (Oxford Nanopore Technologies (ONT), Oxford, UK) using the ARTIC network protocol (accessed the 19th of April; https://artic.network/ncov-2019) Following base calling, data were demultiplexed using ONT Guppy using a high accuracy model. Reads were filtered based on quality and length (400 to 700bp), then mapped to the Wuhan reference genome and primer sites trimmed. Reads were then downsampled to 200x coverage in each direction. Variants were called using nanopolish (https://github.com/jts/nanopolish) and used to determine changes from the reference. Consensus sequences were constructed using reference and variants called.

Pseudotyped Virus Infectivity

VSV System

Plasmids for full-length SARS-Cov-2 Spike were generated from synthetic codon-optimized DNA (Wuhan-Hu-1 isolate, GenBank: MN908947.3) through sub-cloning into the pHCMV3 expression vector, with a stop codon included prior to the HA tag. The D614G variant was generated by site-directed mutagenesis. Positive clones were fully sequenced to ensure that no additional mutations were introduced.

Lentiviruses for stable cell line production were generated by seeding 293T cells at a density of $1x10^6$ cells/well in a 6-well dish. Once the cells reached confluency, they were transfected with 2ug pCaggs-VSV-G, 2ug of lentiviral packaging vector pSPAX2, and 2ug of lentiviral expression plasmid pCW62 encoding ACE2-V5 and the puromycin resistance gene (pCW62-ACE2.V5-PuroR) or TMPRSS2-FLAG and the blasticidin resistance gene (pCW62-TMPRSS2.FLAG-BlastR) using Trans-IT transfection reagent according to manufacturer's instructions. 24 hours post-transfection, media was replaced with fresh DMEM containing 10% FBS and 20mM HEPES. 48 hours post-transfection, supernatants were collected and filtered using a 0.45um syringe filter (VWR Catalog #28200-026).

293T-ACE2 cells were generated by seeding 293T cells at a density of $1x10^6$ cells/well in a 6-well dish. At confluency, cells were transduced with 100uL of ACE2.V5-PuroR lentivirus. 48 hours post-transduction, cells were placed under 5ug/ml puromycin. 293T-ACE2+TMPRSS2 cells were generated by seeding 293T-ACE2 cells at a density of $1x10^6$ cells/well in a 6-well dish. At confluency, cells were transduced with 100uL of TMPRSS2.FLAG-BlastR lentivirus. 48 hours post-transduction, cells were placed under 10ug/ml blasticidin selection.

Recombinant SARS-CoV-2-pseduotyped VSV-ΔG-GFP were generated by transfecting 293T cells with phCMV3 expressing the indicated version of codon-optimized SARS-CoV-2 Spike using TransIT according to the manufacturer's instructions. At 24 hr post-transfection, the medium was removed, and cells were infected with rVSV-G pseudotyped ΔG-GFP parent virus (VSV-G*ΔG-GFP) at MOI = 2 for 2 hours with rocking. The virus was then removed, and the cells were washed twice with OPTI-MEM containing 2% FBS (OPTI-2) before fresh OPTI-2 was added. Supernatants containing rVSV-SARS-2 were removed 24 hours post-infection and clarified by centrifugation.

Viral titrations were performed by seeding cells in 96-well plates at a density sufficient to produce a monolayer at the time of infection. Then, 10-fold serial dilutions of pseudovirus were made and added to cells in triplicate wells. Infection was allowed to proceed for 12-16 hr at 37°C. The cells were then fixed with 4% PFA, washed two times with 1xPBS and stained with Hoescht (1ug/mL in PBS). After two additional washes with PBS, pseudovirus titers were quantified as the number of fluorescent forming units (ffu/mL) using a CellInsight CX5 imager (ThermoScientific) and automated enumeration of cells expressing GFP.

Lentiviral System

Additional assessments of corresponding D614 and G614 Spike pseudotyped viruses were performed by using lentiviral vectors and infection in 293T/ACE2.MF and TZM-bl/ACE2.MF cells (both cell lines kindly provided by Drs. Mike Farzan and Huihui Mu at Scripps). Cells were maintained in DMEM containing 10% FBS, 1% Pen Strep and 3 ug/ml puromycin. An expression plasmid encoding codon-optimized full-length spike of the Wuhan-1 strain (VRC7480), was provided by Drs. Barney Graham and Kizzmekia Corbett at the Vaccine Research Center, National Institutes of Health (USA). The D614G amino acid change was introduced into VRC7480 by site-directed mutagenesis using the QuikChange Lightning Site-Directed Mutagenesis Kit from Agilent Technologies (Catalog # 210518). The mutation was confirmed by full-length spike gene sequencing. Pseudovirions were produced in HEK293T/17 cells (ATCC cat. no. CRL-11268) by transfection using Fugene 6 (Promega Cat#E2692). Pseudovirions for 293T/ACE2 infection were produced by co-transfection with a lentiviral backbone (pCMV ΔR8.2) and firefly luciferase reporter gene (pHR' CMV Luc) (Naldini et al., 1996). Pseudovirions for TZM-bl/ACE2 infection were produced by co-transfection with the Env-deficient lentiviral backbone pSG3ΔEnv (kindly provided by Drs Beatrice Hahn and Feng Gao). Culture supernatants from transfections were clarified of cells by low-speed centrifugation and filtration (0.45 μm filter) and used immediately for infection in 96-well culture plates. 293T/ACE2.MF cells were preseeded at 5,000 cells per well in 96-well black/white culture plates (Perkin-Elmer Catalog # 6005060) one day prior to infection. Sixteen wells were inoculated with 50 ul of a 1:10-dilution of each pseudovirus and incubated for three days. Luminescence was measured using the Promega Luciferase Assay System (Catalog # E1501). For infection of TZM-bl/ACE2.MF cells, 10,000 freshly trypsinized cells were added to 16 wells of a 96-well clear culture plate (Fisher Scientific) and inoculated with undiluted pseudovirus. Luminescence was measured after 2 days in a solid black plate using the Britelite Plus Reporter Gene Assay System (Perkin-Elmer). Luminescence in both assays was measured using a PerkinElmer Life Sciences, Model Victor2 luminometer. HIV-1 p24 content (produced by the backbone vectors) was quantified using the Alliance p24 ELISA Kit (PerkinElmer Health Sciences, Cat# NEK050B001KT). Reported relative luminescence units (RLUs) were adjusted for p24 content.

Neutralization Assay

Pre-titrated amounts of rVSV-SARS-CoV-2 (D614 or G614 variant) were incubated with serially diluted human sera at 37°C for 1 hr before addition to confluent Vero monolayers in 96-well plates. Infection proceeded for 12-16 hr at 37°C in 5% CO_2 before cells were fixed in 4% paraformaldehyde and stained with 1ug/mL Hoescht. Cells were imaged using a CellInsight CX5 imager and infection was quantitated by automated enumeration of total cells and those expressing GFP. Infection was normalized to the percent cells infected with rVSV-SARS-CoV-2 incubated with normal human sera. Data are presented as the relative neutralization for each concentration of sera.

Data Pipeline

Background and General Approach

The Global Initiative for Sharing All Influenza Data (GISAID) (Elbe and Buckland-Merrett, 2017; Shu and McCauley, 2017) has been coordinating SARS-CoV-2 genome sequence submissions and making data available for download since early in the pandemic. At time of this writing, hundreds of sequences were being added every day. These sequences result from extraordinary efforts by a wide variety of institutions and individuals: while an invaluable resource, but are mixed in quality. The complete sequence download includes a large number of partial sequences, with variable coverage, and extensive 'N' runs in many sequences. To assemble a high-quality dataset for mutational analysis, we constructed a data pipeline using some off-the-shelf bioinformatic tools and a small amount of custom code.

From theSARS-CoV-2 sequences available from GISAID, we derived a ''clean'' codon-aligned dataset comprising near-complete viral genomes, without large insertions or deletions (''indels'') or runs of undetermined or ambiguous bases. For convenience in mutation assessment, we generated a codon-based nucleotide multiple sequence alignment, and extracted translations of each reading frame, from which we generated lists of mutations. The cleaning process was in general a process of deletion, with alignment of retained sequences; the following criteria were used to exclude sequences:

1. Fragmented matching (> 20 nt gap in match to reference)
2. Gaps at 5′ or 3′ end (> 3 nt)
3. High numbers of mismatched nucleotides (> 20), 'N' or other ambiguous IUPAC codes.
4. Regions with concentrated ambiguity calls: > 10 in any 50 nt window)

Any sequence matching any of the above criteria was excluded in its entirety.

Sequence Mapping and Alignment

Sequences were mapped to a reference (bases 266:29674 of GenBank entry NC_045512; i.e., the first base of the ORF1ab start codon to the last base of the ORF10 stop codon) using ''nucmer'' from the MUMmer package (version 3.23; Kurtz et al., 2004). The nucmer output ''delta'' file was parsed directly using custom Perl code to partition sequences into the various exclusion categories (Sequence Mapping Table) and to construct a multiple sequence alignment (MSA). The MSA was refined using code derived from the Los Alamos HIV database ''Gene Cutter'' tool code base. At this stage, alignment columns comprising an insertion of a single ''N'' in a single sequence (generating a frameshift) were deleted, and gaps were shifted to conform with codon boundaries.

Using the initial ''good-sequence'' alignment, a low-effort parsimony tree was constructed. Initially, trees were built using PAUP* (Swofford, 2003) with a single replicate heuristic search using stepwise random sequence addition; subsequently, a parsimony ratchet was added; currently, oblong (Goloboff, 2014) is used. Sequences in the alignment were sorted vertically to correspond to the (ladderized) tree, and reference-sequence reading frames were added. See Figure S1 for a pipeline schematic.

Data partitioning and phylogenetic trees

Alignments were made and trees inferred for three distinct data partitions, the longer the alignments, the fewer sequences the sequences (Figure S1.) The full genome tree was used for Figure 7. Trees were inferred by either of two methods: 1. neighbor-joining using a p-distance criterion, (Swofford, 2003) or 2. parsimony heuristic search using a version of the parsimony ratchet (Goloboff, 2014), the general conclusions in Figure 7 were substantiated in both; the parsimony tree is shown.

Global Maps

The Covid-19 pie chart map is generated by overlaying Leaflet (a JavaScript library for interactive maps) pie charts on maps provided by OpenStreetMap. The interface is presented using rocker/shiny, a Docker for Shiny Server.

QUANTIFICATION AND STATISTICAL ANALYSIS

Systematic Regional Analysis of D614/G614 Frequencies

To observe a significant change in the frequency of two SARS-CoV-2 variants in a geographic region, three minimal requirements must be met. Both variants must have been introduced into an area and be co-circulating, data must be sampled for a long enough period to observe a change in frequency, and there must be enough data to be powered adequately to detect a difference.

We use the bioinformatic approaches described above to extract from GISAID all the politically defined geographic regions within the data that met these criteria, to track changes in frequency in a systematic way using all available data. The political/geographical regions we use are strictly hierarchically segmented based on the naming conventions used in GISAID. GISAID data is labeled such that the geographic source is noted first as a continent or Oceania; we call this Level 1. Level 2 is the country of origin of a sample. Level 3 are subcountries and states, and although occasionally level 3 includes a major city in a small country. For this purpose, England, Scotland, and Wales are considered sub-countries of the United Kingdom, and assigned level 3; the sampling in the UK has been the most extensive globally to date. Level 4, is the county or city of origin. The levels are strictly hierarchical, and within a given level, the geographical regions do not overlap. In some cases (e.g., Nepal_Kathmandu and Nepal, Greece_Athens and Greece, Ita-

ly_Veneto_Verona and Italy_Veneto, or Iceland_Reykjavik and Iceland) the sampling in a sub-level exactly matches the sampling in the corresponding upper level, in which case the sub-level is not presented. Levels 3 and 4 are not always available, and the day of sampling is also not always available.

The statistical strategies we use are then applied separately in each country, region or city, and we do not assume that outbreaks in each political subdivision are independent and identically distributed. Instead, our model assumption is that the individuals we test within a region are independent. This assumption may fail if there are sampling biases in a region that change over a given period of time. The G614 form is part of the G clade haplotype that is introduced by travelers, as we discuss in the text, and it is rare for it to arise independently. Our null hypothesis is that the observed shifts in frequency are random nondirectional drift. We have taken two statistical approaches to test this.

Fisher's exact comparison

For this comparison, we used a two-sided Fisher's exact test to compare the G614 and D614 counts in the pre-onset and the post-delay periods, as described in the text, and provides a p value against the null hypothesis that the fraction of D614 and G614 sequences did not change. To be included in the analysis, 15 sequences were required pre-onset, with a mixture of D614 and G614 present such that the rarer form was present at least 3 times; we also required a minimum of 15 sequences be sampled at least 2 weeks later, to create a post-delay set. Only regions for which $p < 0.05$ are considered, based on a two sided-test. We then use a binomial test to evaluate the null hypothesis that in regions where we saw significant change in sampling frequency over time, the shift was as likely to be an increase or a decrease in G614 across geographic regions. This analysis is presented in Figure 1B.

Isotonic Regression

Isotonic regression forms the basis of a one-sided test of the hypothesis for positive selection based on fitting the indicator that the typed strain is G as a logistic regression in which the logarithm of the odds ratio is a non-decreasing function of time. We use the residual deviance of the fitted model as our test statistics. To be included in this analysis, a region was required to have at least 5 sequences each of D614 and G614, and a minimum of 14 sampling days of data available. While we have a composite null hypothesis (the log-odds ratio is non-increasing), assuming that the log-odds ratio remains constant over time leads to tests that have largest power. While the classical chi-square approximation does not hold, we can sample from the constant log-odds ratio by permuting the vector of variant labels, and refitting the isotonic logistic regression. We performed 400 randomizations of the data in each region. Hence the lowest p value we can obtain is 0.0025. The reverse hypothesis, namely than the fraction of G variant decreases with time is also tested by fitting a non-increasing function of time. The isotonic logistic regression was done using R and the cgam package. We applied the bionomial test across regions with a significant change in one direction, as we did for the Fisher's test results. This analysis is presented in Figure 3 and Data S1.

Clinical Data and Modeling

Baseline Comparisons of Clinical Parameters

Univariate analysis showed no associations between the age of individuals and their D614 (median 54.8, IR 39.4-77) or G614 status (median 54.6 (38.7-72.8) (Wilcoxon rank sum $p = 0.37$), nor with D614 and G614 and sex (Fisher's exact $p = 0.32$). Comparing hospitalization and age, the median (IR) are: for all hospitalized, (IP+OCU), 74 years (59-83); for all OP, 44 (32-54), Wilcoxon $p < 2.2e\text{-}16$.

67% of males were hospitalized, versus 33% of females (Fisher's exact p = p value < 2.2e-16).

Modeling PCR Ct

Two PCR Ct methods were used as a surrogate for estimating *in vivo* viral load in the upper respiratory tract, switching methods in mid-April due a shortage of kits. The first method involved nucleic acid extraction; the second method, heat treatment (Fomsgaard and Rosenstierne, 2020).

To assess the impact of available clinical parameters on viral load as measured by PCR Cts, we used a linear model, predicting Ct from PCR method, Sex, Age and D614G variant. This revealed that only the PCR method and the D614G variant were statistically significant. A negative coefficient for the G variant indicated that patients infected by the latter have, on average, a higher viral load, but that that viral load is not impacted by neither age nor sex.

The results from the smaller model are:

Coefficients:

	Estimate	Std. Error	t value	Pr(> |t|)	
(Intercept)	24.301	0.3166	76.757	< 2e-16	***
G614	−0.7763	0.3718	−2.088	0.037	*
Method_2	3.1979	0.3658	8.743	< 2e-16	

Results comparing D614G status for the two methods were also evaluated independently, and the first method showed a significant association between lower Ct values and presence of G614 (Wilcoxon $p = 0.033$), but the second method, with many fewer samples, did not reach significance.

Predicting Hospitalization

The simple Fisher's exact test analysis in Figure 5 indicates that the D614G status is not predictive of hospitalization, even though it is predictive of viral load. We can make a first analysis to predict hospitalization from viral load, gender, age and D614G status:

Coefficients:

	Estimate	Std. Error	z value	Pr(> |z|)	
(Intercept)	−7.548823	0.624270	−12.092	< 2e-16	***
G614	0.112038	0.214107	0.523	0.600779	
Male	1.490789	0.181695	8.205	2.31e-16	***
Age	0.089444	0.005664	15.791	< 2e-16	***
CT	0.069376	0.018243	3.803	0.000143	***
Method_2	−0.358397	0.218856	−1.638	0.101506	

As somewhat expected, the D614G status is not statistically significant, even though viral load is, but the coefficient goes in the opposite direction than we would have intuited: a lower viral load is predictive of a higher probability of hospitalization. Sex (Male) and Age both increase the probability of hospitalization.

Predicting Hospitalization, revisited

Although the above analysis indicates that aa614G does not predict hospitalization directly, it does predict viral load and viral load predicts hospitalization; so there is a concern that aa614G might affect hospitalization, but that this effect is "masked" by the viral load. To explore this hypothesis, we "unmask" the aa614G by using the residuals from the regression of Ct on extraction method and D614G status to get a second predictive model for hospitalization:

Coefficients:

	Estimate	Std. Error	z value	Pr(> |z|)	
(Intercept)	−5.889991	0.393950	−14.951	< 2e-16	***
G614	0.029858	0.209349	0.143	0.886587	
Corrected Ct	0.069276	0.018225	3.801	0.000144	***
Male	1.490690	0.181584	8.209	2.22e-16	***
Age	0.089714	0.005661	15.849	< 2e-16	***

In these regression analyses, the estimated coefficients for age, sex and viral load (corrected or not for method and strain) remain mostly unchanged, and strain still does not have an effect.

All other comparisons were not significant. All coding was done using R. Results of these analysis are presented in the main text and in Figure 5.

Modeling pseudotype virus infectivity

We used a log-normal generalized linear model (GLM) to test whether the G614 variant grew to higher titers than the wild-type D614 virus in Vero, 293T-ACE2 and 293T-ACE2-TMPRSS2 cell lines. The full experiment was repeated twice, each time in triplicate, and the 2 experimental repeats were considered random effects. Viral variant and cell line were considered as fixed effects. On average, across all cell lines, G614 grows to about a 3-fold (2.95) higher titer than D614 ($p = 9x10^{-11}$). A significant interaction was found between viral variant and cell line ($p = 0.002$), indicating that the relative increase of G614 compared to D614 was significantly different across cell lines ($p = 0.002$).

Results of these analysis are presented in Figure 6A.

Sequence quality control

We discovered a sequencing processing error that gave rise to what appeared at first to be a mutation of interest at position 943 (24389 A > C and 24390 C > G) in Spike that was evident in sequences from Belgium. It was frequent enough to be a site of interest, and was tracked. We contacted the group in Belgium, the source of the data, who were already aware of the issue, concurred with our interpretation, and they had been in touch with GISAID with a request to remove the problematic sequences.

We identified the issue with this site as part of another study using a method to detect systematic sequencing errors (Freeman et al., 2020); we are interrogating the quality of available sequencing

data and these positions were highlighted as suspect. We interrogated these positions in the raw sequencing data from Sheffield, and although these two variants are not present in the final consensus sequence from any of the Sheffield isolates, the raw, untrimmed bam files show their presence in only one of the amplicons covering the site (Figure S7A and S7B). We noticed that in fact this position is to the left of the 5′ primer of amplicon 81 in what we believe to be an adaptor sequence. Comparison of the Wuhan reference and the adaptor sequence reveals similarity around this position:

Nanopore adaptor sequence:

CAGCACCTT

The Wuhan reference sequence:

CAGCAAGTT

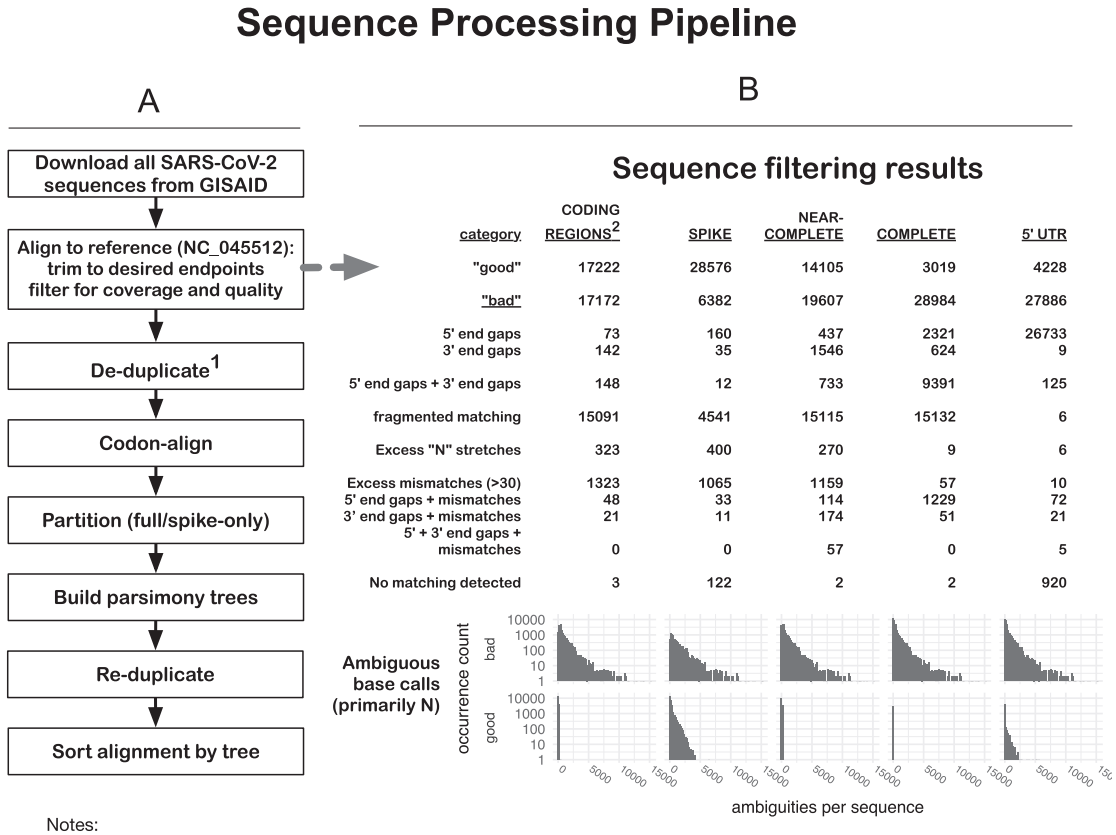
In our validation set, we see a C present at around 50% of called bases at both these positions in raw data but this region is trimmed by the ARTIC pipeline and is therefore not used to call variants and contribute to the final consensus sequence. Although it is evident in amplicon 81, in this region, there is no evidence for these variants in the data from amplicon 80, which also covers these positions. We include a figure (Figure S7) to explain our finding.

In summary this is an error that has arisen due to a combination of improper trimming of adaptor and primer regions from raw sequencing reads before downstream analysis, and the coincidental homology between the nanopore adaptor sequence and the Wuhan reference genome in this region. This is included here as a cautionary note; resolving rare biological mutations and sequencing error will be an important balance going forward in terms of interpretation of rare mutations (De Maio et al., 2020). A recurrent amino acid change like L5F (Figure 7) could potentially result from a recurrent sequencing or sequence processing error (De Maio et al., 2020), or alternatively, it may be of particular interest if it is naturally recurring homoplasy.

ADDITIONAL RESOURCES

Current data updates, analytical results, and webtools: https://cov.lanl.gov

Supplemental Figures



category	CODING REGIONS[2]	SPIKE	NEAR-COMPLETE	COMPLETE	5' UTR
"good"	17222	28576	14105	3019	4228
"bad"	17172	6382	19607	28984	27886
5' end gaps	73	160	437	2321	26733
3' end gaps	142	35	1546	624	9
5' end gaps + 3' end gaps	148	12	733	9391	125
fragmented matching	15091	4541	15115	15132	6
Excess "N" stretches	323	400	270	9	6
Excess mismatches (>30)	1323	1065	1159	57	10
5' end gaps + mismatches	48	33	114	1229	72
3' end gaps + mismatches	21	11	174	51	21
5' + 3' end gaps + mismatches	0	0	57	0	5
No matching detected	3	122	2	2	920

Notes:
1. Multiply occurring identical sequences are reduced to 2 occurrences to so that parsimony-informative sites do not become unique.
2. "Coding regions" subset includes sequences passing error filtering, bounded from the orf1ab start codon to the ORF10 stop codon (NC_045512 genome positions 266-29674).

Figure S1. Schematic of the SARS-CoV-2 Outbreak Sequence Processing Pipeline, Related to Figure 1

The intent of these procedures is to generate, for each of several regions, a set of contiguous codon-aligned sequences, complete in that region, without extensive uncalled bases, large gaps, or regions that are unalignable or highly divergent, in reasonable running time for n > 30,000 ~30kb viral genomes. This allows daily processing of GISAID data to enable us to track mutations. This process provides the foundational data to enable the generation of Figures 1, 2, 3, and 7.

A. Processing procedures:

1. Download all SARS-CoV-2 sequences from GISAID.org (34,607 as of 2020-05-29). The downloaded sequences are stored in compressed form (via bzip2: https://sourceware.org/git/bzip2.git).

2. Align sequences to the SARS-CoV-2 reference sequence (NC_045512), trim to desired endpoints, and filter for coverage and quality. These steps are incorporated in a single Perl script, 'align_to_ref.pl', briefly summarized here: sequences are compressed for identity, then mapped against the given reference sequence using 'nucmer' from the 'MUMmer' package (Kurtz et al., 2004). The nucmer 'delta' file contains locations of matching regions and is parsed and used to, first, partition the sequences into "good" and "bad" subsets, and then to generate alignments from the "good" sequences.

B. Categories of sequences included and excluded from our automated alignments. A series of criteria is used successively to exclude sequences with large internal gaps, excessive five- and three-prime gaps, large numbers of mismatches or ambiguities (> 30) overall, or regions with a high concentration of mismatches or ambiguities (> 10 in any 100 nt subsequence): the counts of these categories of "bad" sequence are shown, for different regional genome alignments. We then create the following different regional subalignments: CODING-REGIONS[2] ("FULL," from the 5′-most start-codon (orf1ab) to the 3′-most stop-codon (ORF10), NC_045512 bases 266-29,674; SPIKE, the complete surface glycoprotein coding region, bases 21,563-25,384; NEAR-COMPLETE ("NEARCOMP," the most-commonly-sequenced region of the genome, bases 55-29,836; COMPLETE, matching the NC_045512 sequence from start up to the poly-A tail, bases 1-29,870; 5′ UTR, the five-prime untranslated region, bases 1-265 only. Generally speaking, the smaller the region, the more sequences are included.
Sequences are trimmed to the extent of the reference (with minimum allowed gaps at 5′ and 3′ ends), following which the pairwise alignments are generated from the matching regions, and a multiple sequence alignment is constructed from the pairwise alignments.

3. De-duplicate[1]. To reduce computational demands, sequences are compressed by identity following trimming to the desired region, by computing a hash value for each sequence (currently the SHA-1 message digest, 160 bits encoded as a 40-character hex string). To prevent the loss of parsimony-informative characters when they occur in identical strings, however, multiple sequences are reduced to a minimum of two occurrences.

(legend continued on next page)

4. Codon-align. Gaps are introduced into the entire compressed alignment so that the alignment column containing the last base of each codon has a number divisible by three; this simplifies processing of translations. Code for this procedure is derived from the GeneCutter tool from the LANL HIV database (https://www.hiv.lanl.gov/content/sequence/GENE_CUTTER/cutter.html).

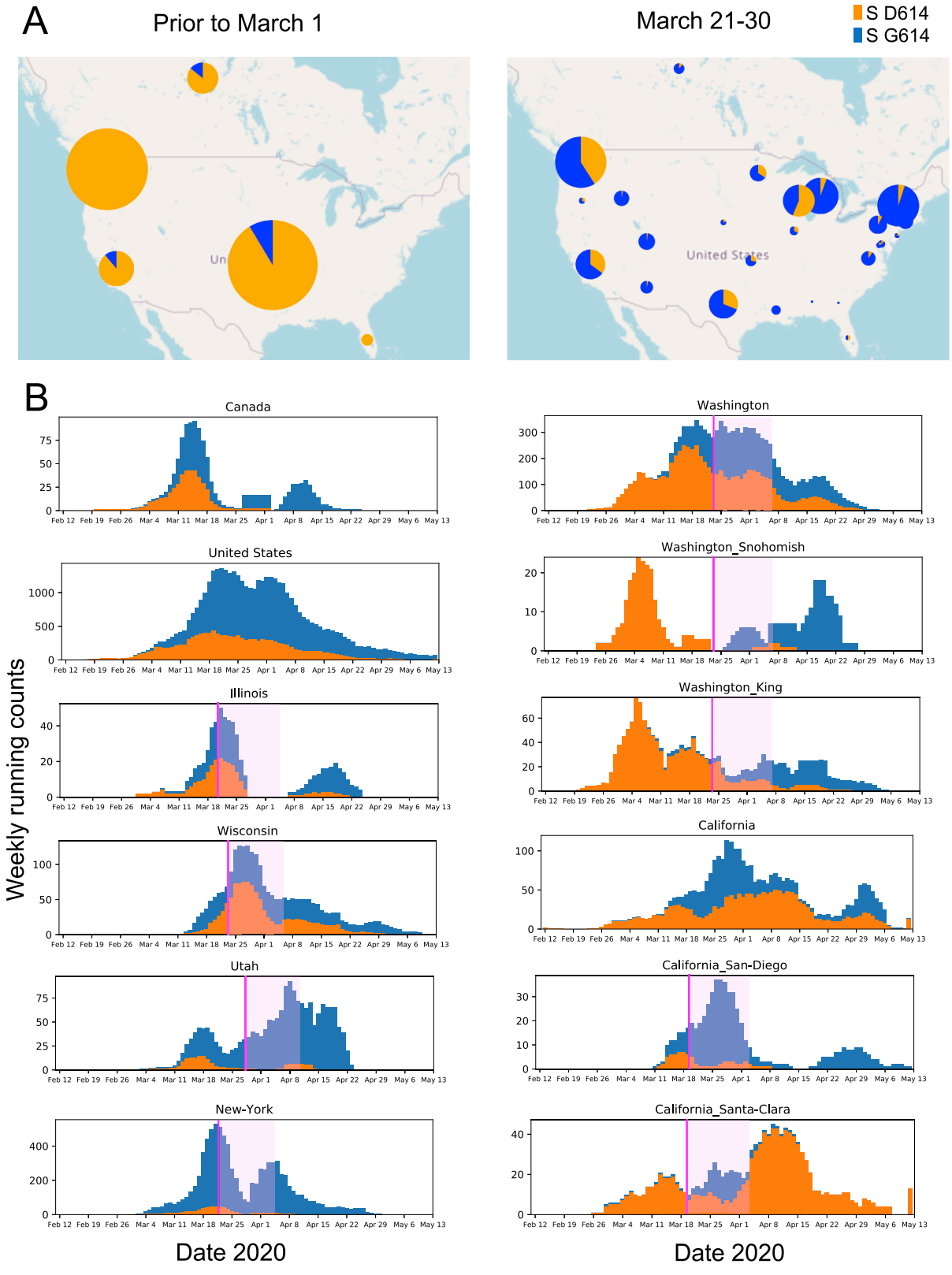
5. Partition (full/spike-only). For subalignments that encompass the spike protein and substantial additional sequence, the spike region is extracted separately, to allow matched comparisons.

6. Build parsimony trees. A brief parsimony search (parsimony ratchet, with 5 replicates) is performed with 'oblong' (Goloboff, 2014) This is intended as an efficient clustering procedure rather than an explicit attempt to achieve an accurate phylogenetic reconstruction, but it appears to yield reasonable results in this situation of a very large number of sequences with a very small number of changes, where more complex models may be subject to overfitting. When multiple most-parsimonious trees are found, only the shortest of these (under a p-distance criterion) is retained. Distance scoring is performed with PAUP* (Swofford, 2003).

7. Re-duplicate (expand, i.e., uncompress). The original sequence names and occurrence counts are restored to FastA format files and the appropriate leaf taxa added to the parsimony trees.

8. Sort alignment by tree. Sequences in the FastA files are sorted by the expanded tree, allowing patterns of mutation to be discerned by inspection.

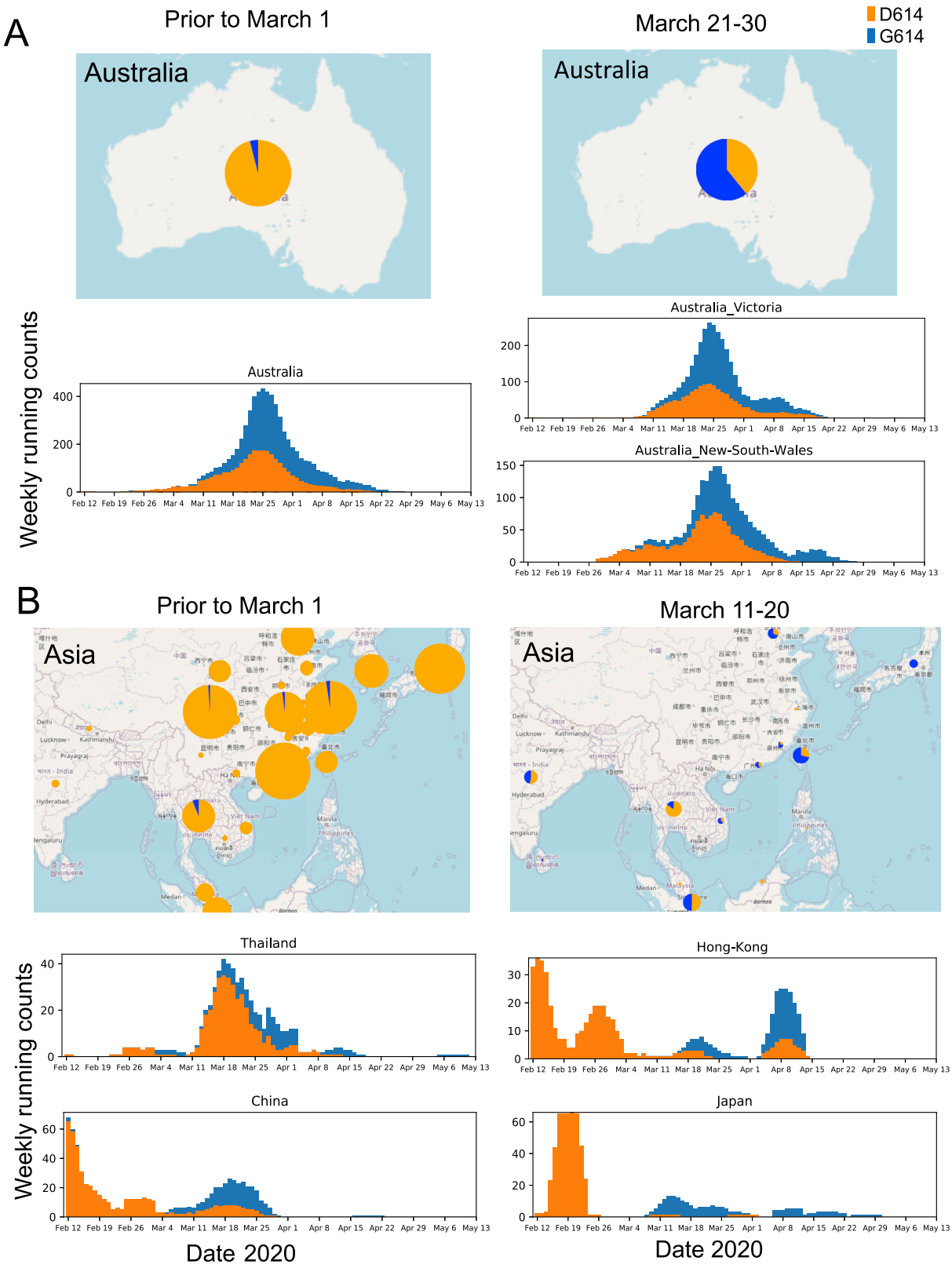
9. Mutations of interest can be readily tracked on the trees to resolve whether they are identified in predominantly in single clades or distributed throughout the tree and likely to be recurring (e.g., Figure 7 (sites of interest with low frequency amino acid substitutions) and Figure S6 (Site 614)).



(legend on next page)

Figure S2. The Increasing Frequency of the D614G Variant over Time in North America, Related to Figure 1

Maps of the relative frequencies of D614 and G614 in North America in two different time windows. B. Weekly running counts of G614 illustrating the timing of its spread in North America. This figure complements Figures 2 and S3, and Figure 1 has details about how to read these figures. When a particular stay-at-home order date was known for a state or county it is shown as a pink line, followed by a light pink block indicating the maximum two-week incubation time. Different counties in California had different stay-at-home order dates (Mar. 16-19) so are not highlighted, but more detail can be seen regarding California in Figure S4. The decline in D614 frequency often continues well after the stay-at-home orders were initiated, and sometimes beyond the 14-day maximum incubation period, when serial reintroduction of the G614 would be unlikely. On the right, Washington State is shown, with details from two heavily sampled counties, Snohomish and King. Both counties had well-established ongoing D614 epidemics when G614 variants were introduced, undoubtedly by travelers. Washington state's stay-at-home order was initiated March 24. At this time there were 1170 confirmed cases in King County, and 614 confirmed cases in Snohomish County. (Confirmed COVID19 case count data from: COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University). Testing was limited, and so this is lower bound on actual cases. Of the sequences sampled by March 24, 95% from King County (153/161) and 100% from Snohomish County (33/33) were the original D614 form (Part B, details at https://cov.lanl.gov/). By mid-April, D614 was rarely sampled. Whatever the geographic origin of the G614 variants that entered these counties, and whether one or if multiple G614 variants were introduced, the rapid expansion of G614 variants occurred in the framework of well-established local D614 variant epidemics. Santa Clara county is one of the two exceptions to the pattern of D614 decline in Figure 1B: details are provided in Figures S4A and S4C.



(legend on next page)

Figure S3. The Increasing Frequency of the D614G Variant over Time in Australia and Asia, Related to Figures 1 and 2

This figure complements Figure 2 and Figure S2, and Figure 1 has details about how to read these figures. The plot representing national sampling in Australia is on the left, with two regional subsets of the data on the right. In each case a local epidemic started with the D614 variant, and despite being well established, the G614 variant soon dominates the sampling. Only limited recent sampling from Asia is currently available in GISAID; to include more samples on the map the 10-day period between March 11-20, is shown rather than the period between March 21-30; even the limited sampling mid-March the supports the repeated pattern of a shift to G614. The Asian epidemic was overwhelmingly D614 through February, and despite this, G614 repeatedly becomes prominent in sampling by mid-March.

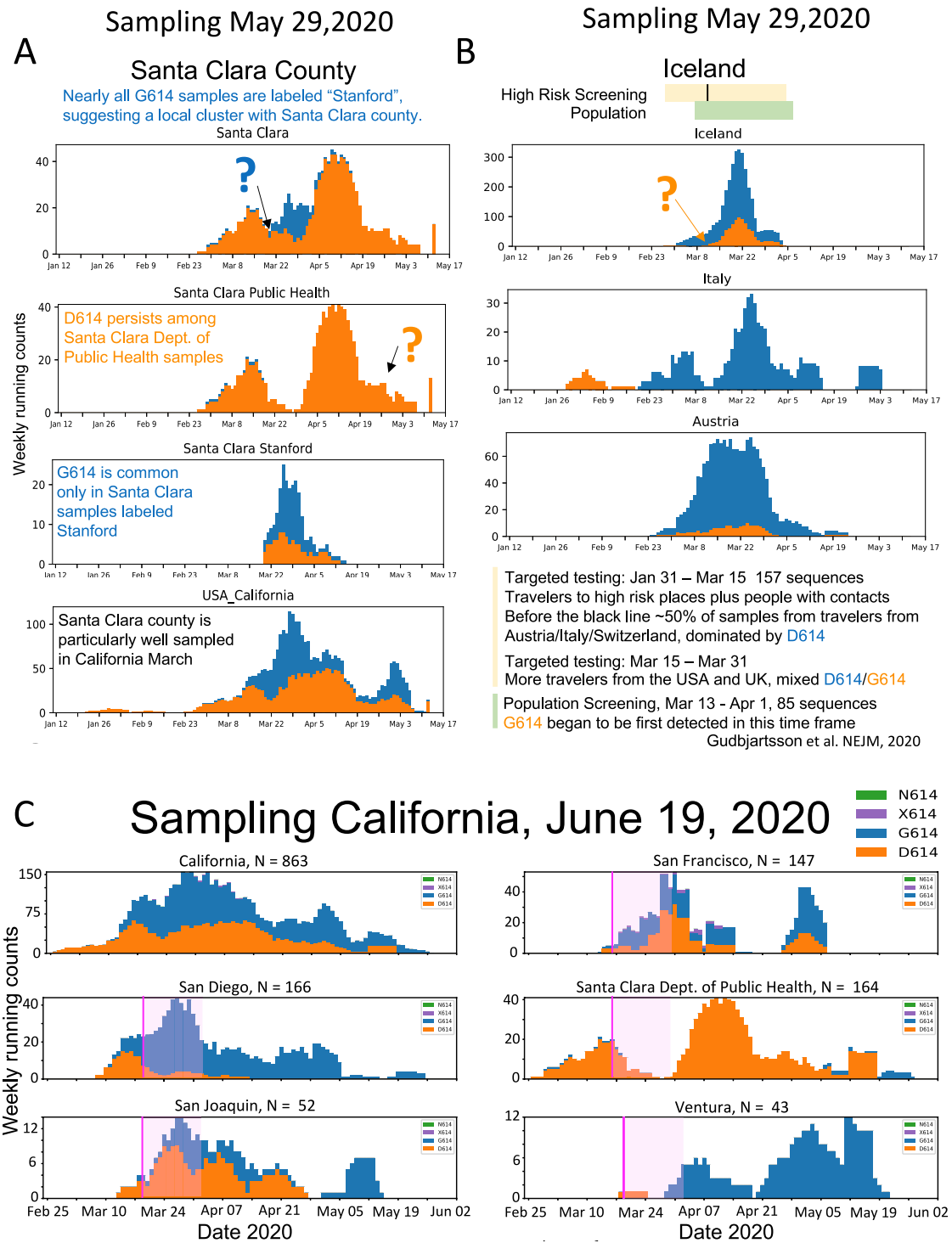
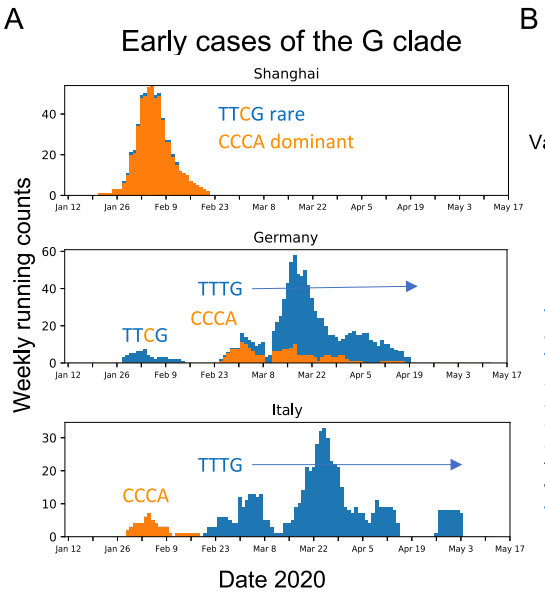


Figure S4. Two Exceptions to the Pattern of Increasing Frequency of the G614 Variant over Time, from Figure 1B, Related to Figure 1

A. Details regarding Santa Clara county, the only exceptional pattern at the county/city level in Figure 1B. Many samples from the Santa Clara County Department of Public Heath (DPH) were obtained from March into May, and D614 has steadily dominated the local epidemic among those samples. The subset of Santa Clara county samples specifically labeled "Stanford," however, were sampled over a few weeks mid-March through early April, and have a mixture of both the G614 and D614 forms. These distinct patterns suggest relatively little mixing between the two local epidemics. Why Santa Clara county DPH samples should maintain the original form is unknown, but one possibility is that they may represent a relatively isolated community that had limited exposure to the G614 form, and G614 may not have had the opportunity to become established in this community – though this may be changing, see Part C. The local stay-at-home orders were

(legend continued on next page)

initiated relatively early, March 16, 2020. B. Details regarding Iceland, the only country with an exceptional pattern from Figure 1B. All Icelandic samples are from Reykjavik, and only G614 variants were initially observed there, with a modest but stable introduction of the original form D614 in mid-March. This atypical pattern might be explained by local sampling. The Icelanders conducted a detailed study of their early epidemic (Gudbjartsson et al., 2020), and all early March samples were collected from high risk travelers from Europe and people in contact with people who were ill; the majority of the traveler samples from early March were from people coming in from Italy and Austria, and G614 dominated both regions. On March 13, they began to sequence samples from local population screening, and on March 15, more travelers from the UK and USA with mixed G614/D614 infections began to be sampled in the high-risk group, and those events were coincident with the appearance of D614. C. Updated data regarding California from the June 19, 2020 GISAID sampling. Most of the analysis in this paper was undertaken using the May 29, 2020 GISAID download, but as California was an interesting outlier, and more recent sampling conducted while the paper was under review was informative, we have included some additional plots from California data that were available at the time of our final response to review, on June 19th. Informative examples from well-sampled local regions are shown. Stay-at-home order dates are shown as a pink line, followed by a light pink block indicating the maximum two-week incubation time. N indicates the number of available sequences. Overall California, and specifically, San Diego and San Joaquin, show a clear shift from D614 to G614. The transition for San Joaquin was well after the stay-at-home orders and incubation period had passed. San Francisco shows a trend toward G614. Santa Clara DPH, which was essentially all D614 in our May 29th GISAID download, had 7 G614 forms sampled in late May that were evident in our June 19th GISAID download. Ventura is an example of a setting that was essentially all G614 when it began to be sampled significantly in early April, so a transition cannot be tracked; i.e., we cannot differentiate in such cases whether the local epidemic originated as a G614 epidemic, or whether it went through a transition from D614 to G614 prior to sampling. The figures in Parts A, B, and C can be recreated with more current data at http://cov.lanl.govcontent/index.



B

G-clade mutations (C3037T, C14408T, A23403G) CCA -> TTG
Plus the linked mutation in the UTR: C241T CCCA -> TTTG

11805	TTG	(72.03%)	9692	TTTG	(71.65%)
4582	CCA	(27.96%)	3835	CCCA	(28.35%)

Variants:

53	CTG	51	TCTG	5	CTCG
39	TCG	32	TTCG	4	CCTA
16	CCG	13	CTTG	3	TCTA
9	TTA	11	TCCA	2	CTTA
8	CTA	9	TCCG	2	CTCA
5	TCA	7	CCCG	1	TTCA
1	ACA	6	TTTA	1	CCTG

Earliest examples in GISAID:
TTCG: Germany, Jan 2020: cluster of cases late Jan.-Feb.
One example: Germany/BavPat1/EPI_ISL_406862|2020-01-28
TTCG: Sampled several times in China, *e.g.*:
Sichuan/SC-PHCC1-022/EPI_ISL_451345|2020-01-24
Shanghai/SH0025/EPI_ISL_416334|2020-02-06
Guangzhou/GZMU0019/EPI_ISL_429080|2020-02-05
CCCG: Sampled twice in early Feb., Wuhan and Thailand
Thailand/Samut_prakarn_840/EPI_ISL_447919|2020-02-04
Wuhan/HBCDC-HB-06/EPI_ISL_412982|2020-02-07
TTTG: First identified in Italy; within 10 days sampled in many in countries in Europe, the USA, Mexico
First sample: Italy/CDG1/2020|EPI_ISL_412973|2020-02-20

Figure S5. Relationships among the Earliest Examples of the G Clade and Other Early Epidemic Samples, Related to Figures 1 and 2

Weekly running counts of the earliest forms of the virus carrying G614. B. The GISAID G clade is based on a 4 base haplotype that distinguishes it from the original Wuhan form. Part B shows the tallies of all of the variants among the 4 base mutations that are the foundation of the G clade haplotype, versus the bases found in the original Wuhan reference strain, and highlights of some of the earliest identified sequences bearing these mutations. We first consider just the 3 mutations that are in coding regions, to enable using an alignment that contains a larger set of sequences. One is in the RdRp protein (nucleotide C14408T resulting in a P323L amino acid change), one in Spike (nucleotide A23403G resulting in the D614G amino acid change) and one is silent (C3037T). The other mutation is in the 5′ UTR (C241T), and tallies based on all 4 positions are done separately, as they are based on an alignment with fewer sequences. The earliest examples of a partial haplotype, TTCG, are found in Germany and China (Parts A and B). This form was present in Shanghai but never expanded (Part A). A cluster of infections that all carried this form were identified in Germany, but they did not expand and were subsequently replaced with the original Wuhan form, only to be replaced again by sequences that carry the full 4 base haplotype variant TTTG. The first example in our alignments (see Figure S1) found to carry the full 4 base haplotype was sampled in Italy on Feb. 20. The first cases in Italy were of the original Wuhan form CCCA form, but by the end of February TTTG was the only form sampled in Italy, and it is the TTTG form that has come the dominate the pandemic. Of note, the TTCG form did not expand, and it lacked the RdRp P323L change, raising the possibility that the P323L change may contribute to a selective advantage of the haplotype. TTTG and CCCA are almost always linked in SARS-CoV-2 genomes, > 99.9% of the time (Part B). The number of cases where the clade haplotype is disrupted, and the form of the disruption, are all noted in part B (not including ambiguous base calls). Some cases of a disrupted haplotype may be due to recombination events and not *de novo* mutations. Given the fact that these two forms are were co-circulating in many communities throughout the spring of 2020, and the fact that disruptions in the 4 base pattern are rare, suggests that recombination is overall relatively rare among pandemic sequences.

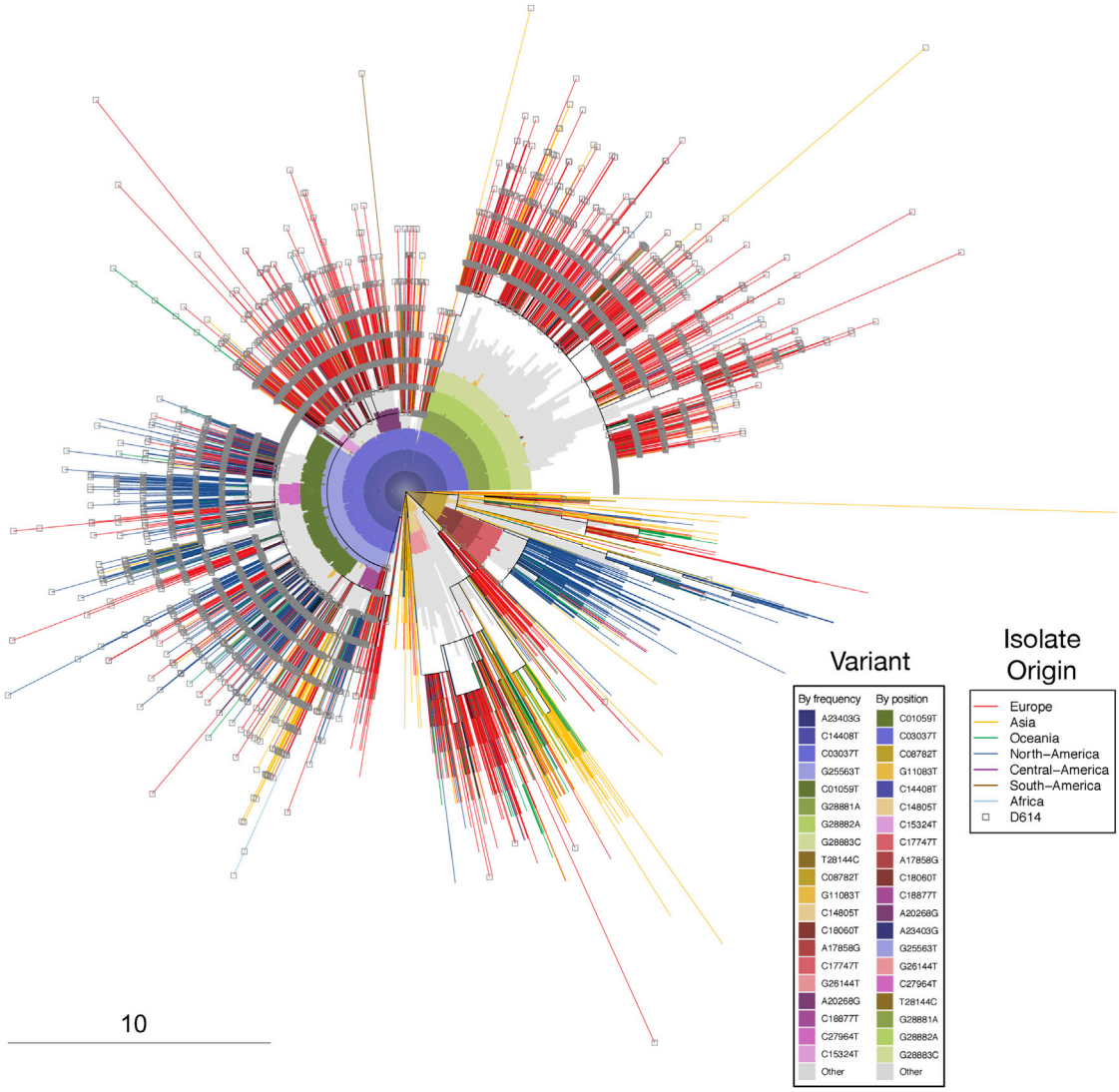
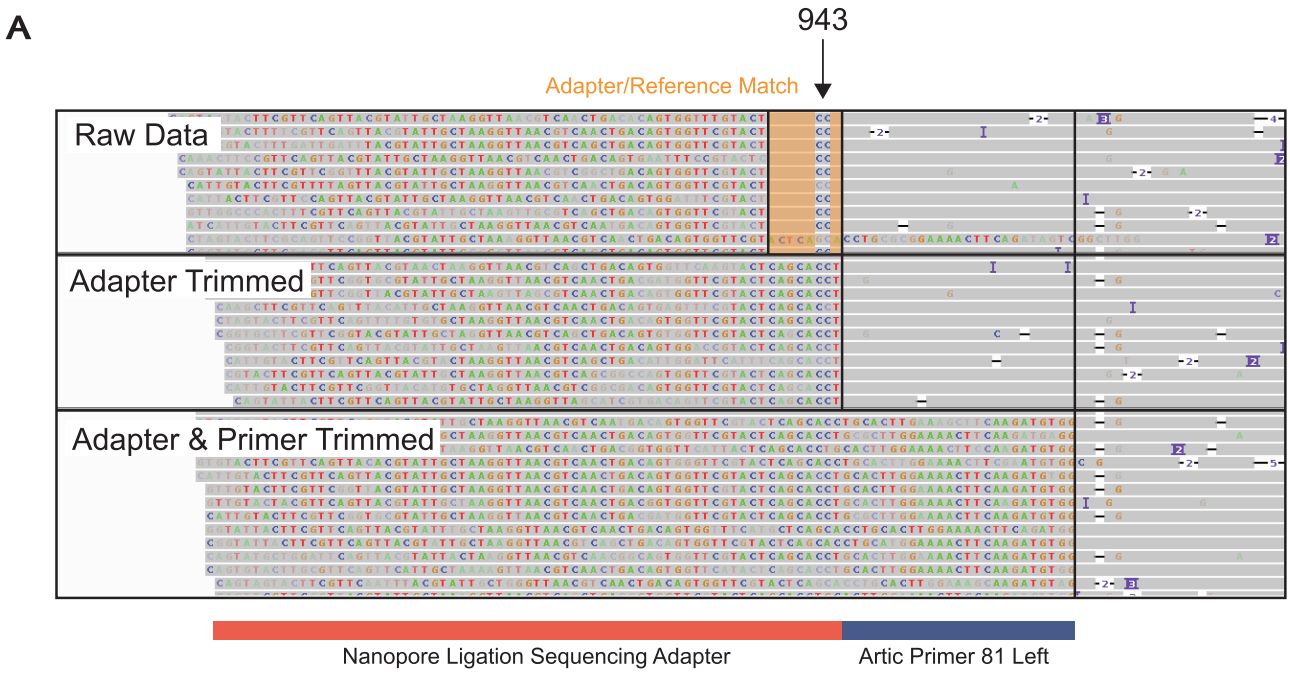


Figure S6. Distribution of A23403G (D614G) Mutation and Other Mutations on an Approximate Phylogenetic Tree Using Parsimony, Related to Figure 7

This tree the same as the tree shown in Figure 6A, but highlights complementary information: the G614 substitution, and patterns of bases that underly the clades. It is based on the "FULL" alignment of 17,760 sequences, from the June 2 alignment, described in the pipeline in Figure S1, from the beginning of (orf1ab) to the last stop-codon (ORF10), NC_045512 bases 266-29,674). The outer element is a radial presentation of a full-coding-region parsimony tree; branches are colored by the global region of origin for each virus isolate. The inner element is a radial bar chart showing the identity of common mutations (any of the top 20 single-nucleotide mutations from the June 2 alignment), so that sectors of the tree containing a particular mutation at high frequency are subtended by an inner colored arc; mutations not in the top 20 are presented together in gray. The tree is rooted on a reference sequence derived from the original Wuhan isolates (GenBank accession number NC_045512), at the 3 o'clock position. Branch ends representing sequence isolates bearing the D614G change are decorated with a gray square; sectors of the tree containing that mutation are subtended by a dark blue arc in the inner element; other mutations are denoted by different colors. As an example, in this tree, the region from approximately 12:30 to 3 o'clock represents GISAID's "GR" clade, defined both by mutations we are tracking in this paper that carry the G614 variant (the GISAID G clade, defined by mutations A23403G, C14408T, C3037T, and a mutation in the 5′ UTR (C241T, not shown here), and an additional 3-position polymorphism: G28881A + G28882A + G28883C. These base substitutions are contiguous and result two amino acid changes, including N-G204R, hence GISAID's "GR clade" name. Close examination of this triplet in sequences from the Sheffield dataset suggests the mutations are not a sequencing artifact. The outer phylogenetic tree was computed using oblong (see STAR Methods), and plotted with the APE package in R. The inner element is a bar chart plotted with polar coordinates using the gglot2 package in R. The frequency of the GR clade appears to be increased in the UK and Europe as a subset of the regional G clade expansion, given that both carry G614D.



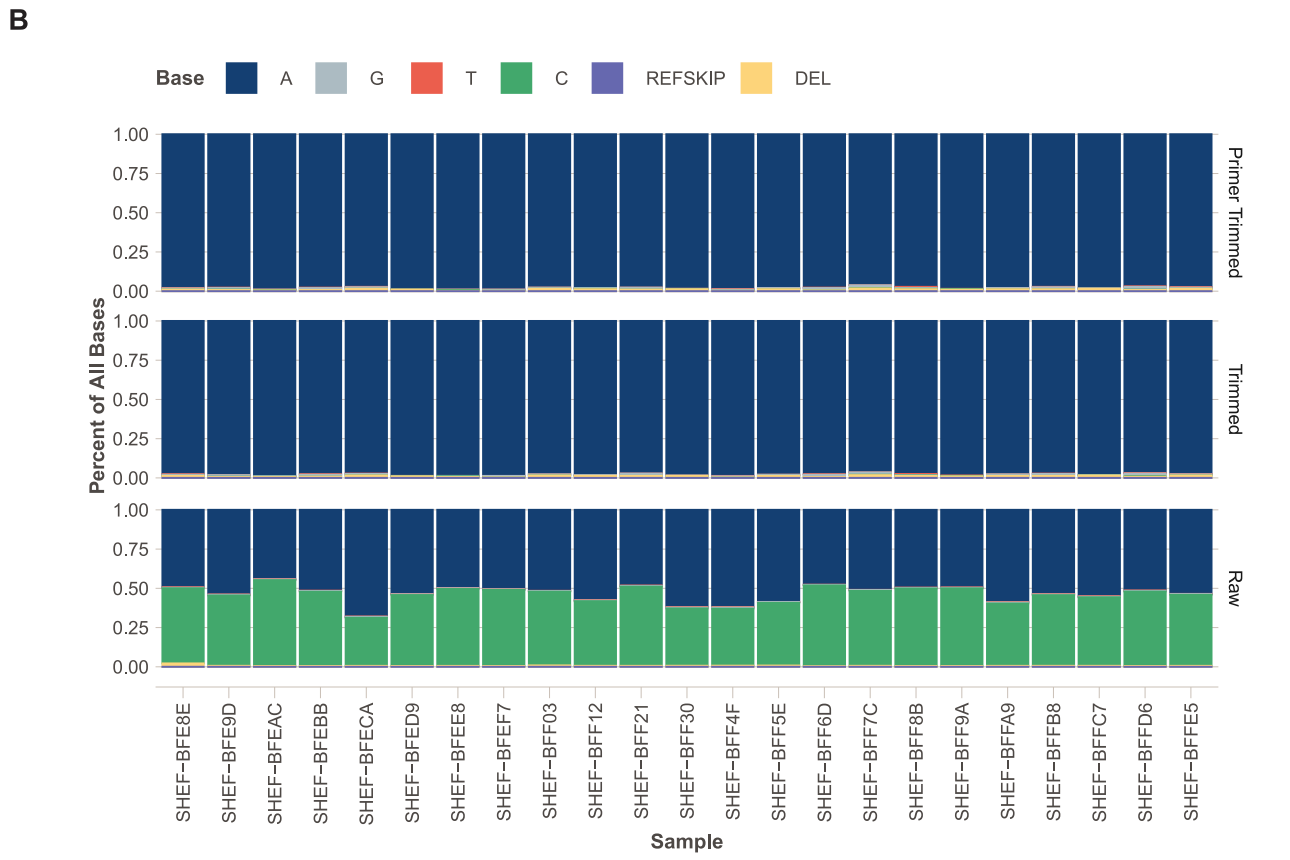


Figure S7. Investigation of S943P, Related to Figure 7

A. IGV plots showing bam files from nanopore sequencing data of amplicons produced by the Artic network protocol. Raw data from amplicon 81 contains a portion of adaptor sequence which is homologous to the reference genome, apart from the C variants which lead to a S943P mutation call. This region is therefore included in variant calling if location-based trimming is not carried out. Subsequent panels show that this region is soft clipped when trimming adapters and primers and is therefore not available for variant calling. B. Base frequencies at position 24389 in 23 samples from the Sheffield data show that C is present in half of the reads in the raw data, but is absent from trimmed and primer trimmed data. This figure is also associated with the STAR Methods.

THE Hospitalist

shm

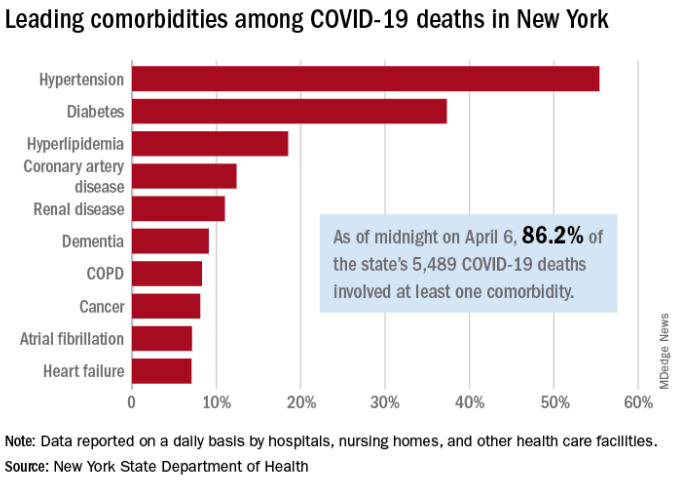
ALL CONTENT

Comorbidities the rule in New York's COVID-19 deaths

Publish date: April 8, 2020

Author(s): Richard Franki

In New York state, just over 86% of reported COVID-19 deaths involved at least one comorbidity, according to the state's department of health.



As of midnight on April 6, there had been 5,489 fatalities caused by COVID-19 in the state, of which 86.2% (4,732) had at least one underlying condition, the New York State Department of Health reported April 7 on its COVID-19 tracker <https://covid19tracker.health.ny.gov/views/NYS-COVID19-Tracker/NYSDOHCOVID-19Tracker-Fatalities?%3Aembed=yes&%3Atoolbar=no&%3Atabs=n> .

The leading comorbidity, seen in 55.4% of all deaths, was hypertension. In comparison, a recent estimate <https://millionhearts.hhs.gov/data-reports/hypertension-prevalence.html> from the U.S. Department of Health

& Human Services put the prevalence of high blood pressure at about 45% in the overall adult population.

In New York, the rest of the 10 most common comorbidities in COVID-19 fatalities were diabetes (37.3%), hyperlipidemia (18.5%), coronary artery disease (12.4%), renal disease (11.0%), dementia (9.1%), chronic obstructive pulmonary disease (8.3%), cancer (8.1%), atrial fibrillation (7.1%), and heart failure (7.1%), the NYSDOH said.

RELATED
Comorbidities more common in hospitalized COVID-19 patients

Other data on the tracker site show that 63% of all deaths involved a patient who was aged 70 years or older and that 61% of COVID-19 patients who have died in New York were male and 38.8% were female (sex unknown for 0.2%). Among all individuals who have tested positive, 54.8% were male and 44.6% were female (sex unknown for 0.6%).

As of the end of day on April 6, a total of 340,058 persons had been tested in the state and 40.8% (138,863) were positive for the SARS-CoV-2 virus. By county, the highest positive rates are in New York City: Queens at 57.4%, Brooklyn at 52.4%, and the Bronx at 52.3%, according to the NYSDOH.

rfranki@mdedge.com

JAMA | Original Investigation

Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area

Safiya Richardson, MD, MPH; Jamie S. Hirsch, MD, MA, MSB; Mangala Narasimhan, DO; James M. Crawford, MD, PhD; Thomas McGinn, MD, MPH; Karina W. Davidson, PhD, MASc; and the Northwell COVID-19 Research Consortium

IMPORTANCE There is limited information describing the presenting characteristics and outcomes of US patients requiring hospitalization for coronavirus disease 2019 (COVID-19).

OBJECTIVE To describe the clinical characteristics and outcomes of patients with COVID-19 hospitalized in a US health care system.

DESIGN, SETTING, AND PARTICIPANTS Case series of patients with COVID-19 admitted to 12 hospitals in New York City, Long Island, and Westchester County, New York, within the Northwell Health system. The study included all sequentially hospitalized patients between March 1, 2020, and April 4, 2020, inclusive of these dates.

EXPOSURES Confirmed severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection by positive result on polymerase chain reaction testing of a nasopharyngeal sample among patients requiring admission.

MAIN OUTCOMES AND MEASURES Clinical outcomes during hospitalization, such as invasive mechanical ventilation, kidney replacement therapy, and death. Demographics, baseline comorbidities, presenting vital signs, and test results were also collected.

RESULTS A total of 5700 patients were included (median age, 63 years [interquartile range {IQR}, 52-75; range, 0-107 years]; 39.7% female). The most common comorbidities were hypertension (3026; 56.6%), obesity (1737; 41.7%), and diabetes (1808; 33.8%). At triage, 30.7% of patients were febrile, 17.3% had a respiratory rate greater than 24 breaths/minute, and 27.8% received supplemental oxygen. The rate of respiratory virus co-infection was 2.1%. Outcomes were assessed for 2634 patients who were discharged or had died at the study end point. During hospitalization, 373 patients (14.2%) (median age, 68 years [IQR, 56-78]; 33.5% female) were treated in the intensive care unit care, 320 (12.2%) received invasive mechanical ventilation, 81 (3.2%) were treated with kidney replacement therapy, and 553 (21%) died. As of April 4, 2020, for patients requiring mechanical ventilation (n = 1151, 20.2%), 38 (3.3%) were discharged alive, 282 (24.5%) died, and 831 (72.2%) remained in hospital. The median postdischarge follow-up time was 4.4 days (IQR, 2.2-9.3). A total of 45 patients (2.2%) were readmitted during the study period. The median time to readmission was 3 days (IQR, 1.0-4.5) for readmitted patients. Among the 3066 patients who remained hospitalized at the final study follow-up date (median age, 65 years [IQR, 54-75]), the median follow-up at time of censoring was 4.5 days (IQR, 2.4-8.1).

CONCLUSIONS AND RELEVANCE This case series provides characteristics and early outcomes of sequentially hospitalized patients with confirmed COVID-19 in the New York City area.

Audio and Video and Supplemental content

Related article at jamahealthforum.com

Author Affiliations: Institute of Health Innovations and Outcomes Research, Feinstein Institutes for Medical Research, Northwell Health, Manhasset, New York (Richardson, Hirsch, McGinn, Davidson); Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Northwell Health, Hempstead, New York (Richardson, Hirsch, Narasimhan, Crawford, McGinn, Davidson); Department of Information Services, Northwell Health, New Hyde Park, New York (Hirsch).

Group Information: The Northwell COVID-19 Research Consortium authors and investigators appear at the end of the article.

Corresponding Author: Karina W. Davidson, PhD, Northwell Health, 130 E 59th St, Ste 14C, New York, NY 10022 (KDavidson2@northwell.edu).

JAMA. 2020;323(20):2052-2059. doi:10.1001/jama.2020.6775
Published online April 22, 2020. Corrected on April 24, 2020.

The first confirmed case of coronavirus disease 2019 (COVID-19) in the US was reported from Washington State on January 31, 2020.[1] Soon after, Washington and California reported outbreaks, and cases in the US have now exceeded total cases reported in both Italy and China.[2] The rate of infections in New York, with its high population density, has exceeded every other state, and, as of April 20, 2020, it has more than 30% of all of the US cases.[3]

Limited information has been available to describe the presenting characteristics and outcomes of US patients requiring hospitalization with this illness. In a retrospective cohort study from China, hospitalized patients were predominantly men with a median age of 56 years; 26% required intensive care unit (ICU) care, and there was a 28% mortality rate.[4] However, there are significant differences between China and the US in population demographics,[5] smoking rates,[6] and prevalence of comorbidities.[7]

This study describes the demographics, baseline comorbidities, presenting clinical tests, and outcomes of the first sequentially hospitalized patients with COVID-19 from an academic health care system in New York.

Key Points

Question What are the characteristics, clinical presentation, and outcomes of patients hospitalized with coronavirus disease 2019 (COVID-19) in the US?

Findings In this case series that included 5700 patients hospitalized with COVID-19 in the New York City area, the most common comorbidities were hypertension, obesity, and diabetes. Among patients who were discharged or died (n = 2634), 14.2% were treated in the intensive care unit, 12.2% received invasive mechanical ventilation, 3.2% were treated with kidney replacement therapy, and 21% died.

Meaning This study provides characteristics and early outcomes of patients hospitalized with COVID-19 in the New York City area.

Methods

The study was conducted at hospitals in Northwell Health, the largest academic health system in New York, serving approximately 11 million persons in Long Island, Westchester County, and New York City. The Northwell Health institutional review board approved this case series as minimal-risk research using data collected for routine clinical practice and waived the requirement for informed consent. All consecutive patients who were sufficiently medically ill to require hospital admission with confirmed severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection by positive result on polymerase chain reaction testing of a nasopharyngeal sample were included. Patients were admitted to any of 12 Northwell Health acute care hospitals between March 1, 2020, and April 4, 2020, inclusive of those dates. Clinical outcomes were monitored until April 4, 2020, the final date of follow-up.

Data were collected from the enterprise electronic health record (Sunrise Clinical Manager; Allscripts) reporting database, and all analyses were performed using version 3.5.2 of the R programming language (R Project for Statistical Computing; R Foundation). Patients were considered to have confirmed infection if the initial test result was positive or if it was negative but repeat testing was positive. Repeat tests were performed on inpatients during hospitalization shortly after initial test results were available if there was a high clinical pretest probability of COVID-19 or if the initial negative test result had been judged likely to be a false-negative due to poor sample collection. Transfers from one in-system hospital to another were merged and considered as a single visit. There were no transfers into or out of the system. For patients with a readmission during the study period, data from the first admission are presented.

Data collected included patient demographic information, comorbidities, home medications, triage vitals, initial laboratory tests, initial electrocardiogram results, diagnoses during the hospital course, inpatient medications, treatments (including invasive mechanical ventilation and kidney replacement therapy), and outcomes (including length of stay, discharge, readmission, and mortality). Demographics, baseline comorbidities, and presenting clinical studies were available for all admitted patients. All clinical outcomes are presented for patients who completed their hospital course at study end (discharged alive or dead). Clinical outcomes available for those in hospital at the study end point are presented, including invasive mechanical ventilation, ICU care, kidney replacement therapy, and length of stay in hospital. Outcomes such as discharge disposition and readmission were not available for patients in hospital at study end because they had not completed their hospital course. Home medications were reported based on the admission medication reconciliation by the inpatient-accepting physician because this is the most reliable record of home medications. Final reconciliation has been delayed until discharge during the current pandemic. Home medications are therefore presented only for patients who have completed their hospital course to ensure accuracy.

Race and ethnicity data were collected by self-report in prespecified fixed categories. These data were included as study variables to characterize admitted patients. Initial laboratory testing was defined as the first test results available, typically within 24 hours of admission. For initial laboratory testing and clinical studies for which not all patients had values, percentages of total patients with completed tests are shown. The Charlson Comorbidity Index predicts 10-year survival in patients with multiple comorbidities and was used as a measure of total comorbidity burden.[8] The lowest score of 0 corresponds to a 98% estimated 10-year survival rate. Increasing age in decades older than age 50 years and comorbidities, including congestive heart disease and cancer, increase the total score and decrease the estimated 10-year survival. A total of 16 comorbidities are included. A score of 7 points and above corresponds to a 0% estimated 10-year survival rate. Acute kidney injury was identified as an increase in serum creatinine by 0.3 mg/dL or more (≥26.5 μmol/L) within 48 hours or an increase in

Table 1. Baseline Characteristics of Patients Hospitalized With COVID-19

	No. (%)
Demographic information	
Total No.	5700
Age, median (IQR) [range], y	63 (52-75) [0-107]
Sex	
Female	2263 (39.7)
Male	3437 (60.3)
Race[a]	
No.	5441
African American	1230 (22.6)
Asian	473 (8.7)
White	2164 (39.8)
Other/multiracial	1574 (28.9)
Ethnicity[a]	
No.	5341
Hispanic	1230 (23)
Non-Hispanic	4111 (77)
Preferred language non-English	1054 (18.5)
Insurance	
Commercial	1885 (33.1)
Medicaid	1210 (21.2)
Medicare	2415 (42.4)
Self-pay	95 (1.7)
Other[b]	95 (1.7)
Comorbidities	
Total No.	5700
Cancer	320 (6)
Cardiovascular disease	
Hypertension	3026 (56.6)
Coronary artery disease	595 (11.1)
Congestive heart failure	371 (6.9)
Chronic respiratory disease	
Asthma	479 (9)
Chronic obstructive pulmonary disease	287 (5.4)
Obstructive sleep apnea	154 (2.9)
Immunosuppression	
HIV	43 (0.8)
History of solid organ transplant	55 (1)
Kidney disease	
Chronic[c]	268 (5)
End-stage[d]	186 (3.5)
Liver disease	
Cirrhosis	19 (0.4)
Chronic	
Hepatitis B	8 (0.1)
Hepatitis C	3 (0.1)
Metabolic disease	
Obesity (BMI ≥30)	1737 (41.7)
No.	4170
Morbid obesity (BMI ≥35)	791 (19.0)
No.	4170
Diabetes[e]	1808 (33.8)
Never smoker	3009 (84.4)
No.	3567
Comorbidities[f]	
None	350 (6.1)
1	359 (6.3)
>1	4991 (88)
Total, median (IQR)	4 (2-8)
Charlson Comorbidity Index score, median (IQR)[g]	4 (2-6)

(continued)

Table 1. Baseline Characteristics of Patients Hospitalized With COVID-19 (continued)

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); COVID-19, coronavirus disease 2019; IQR, interquartile range.

[a] Race and ethnicity data were collected by self-report in prespecified fixed categories.

[b] Other insurance includes military, union, and workers' compensation.

[c] Assessed based on a diagnosis of chronic kidney disease in medical history by *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* (*ICD-10*) coding.

[d] Assessed based on a diagnosis of end-stage kidney disease in medical history by *ICD-10* coding.

[e] Assessed based on a diagnosis of diabetes mellitus and includes diet-controlled and non–insulin-dependent diabetes.

[f] Comorbidities listed here are defined as medical diagnoses included in medical history by *ICD-10* coding. These include, but are not limited to, those presented in the table.

[g] Charlson Comorbidity Index predicts the 10-year mortality for a patient based on age and a number of serious comorbid conditions, such as congestive heart failure or cancer. Scores are summed to provide a total score to predict mortality. The median score of 4 corresponds to a 53% estimated 10-year survival and reflects a significant comorbidity burden for these patients.

serum creatinine to 1.5 times or more baseline within the prior 7 days compared with the preceding 1 year of data in acute care medical records. This was based on the Kidney Disease: Improving Global Outcomes (KDIGO) definition.[9] Acute hepatic injury was defined as an elevation in aspartate aminotransferase or alanine aminotransferase of more than 15 times the upper limit of normal.

Results

A total of 5700 patients were included (median age, 63 years [interquartile range {IQR}, 52-75; range, 0-107 years]; 39.7% female) (**Table 1**). The median time to obtain polymerase chain reaction testing results was 15.4 hours (IQR, 7.8-24.3). The most common comorbidities were hypertension (3026, 56.6%), obesity (1737, 41.7%), and diabetes (1808, 33.8%). The median score on the Charlson Comorbidity Index was 4 points (IQR, 2-6), which corresponds to a 53% estimated 10-year survival and reflects a significant comorbidity burden for these patients. At triage, 1734 patients (30.7%) were febrile, 986 (17.3%) had a respiratory rate greater than 24 breaths/minute, and 1584 (27.8%) received supplemental oxygen (**Table 2** and **Table 3**). The first test for COVID-19 was positive in 5517 patients (96.8%), while 183 patients (3.2%) had a negative first test and positive repeat test. The rate of

Table 2. Presentation Vitals and Laboratory Results of Patients Hospitalized With COVID-19

Triage vitals[a]	No. (%)	No.	Reference ranges
Temperature >38 °C	1734 (30.7)	5644	
Temperature, median (IQR), °C	37.5 (36.9-38.3)		
Oxygen saturation			
<90%	1162 (20.4)	5693	
% Median (IQR)	95 (91-97)		
Received supplemental oxygen at triage	1584 (27.8)	5693	
Respiratory rate >24 breaths/min	986 (17.3)	5695	
Heart rate			
≥100 beats/min	2457 (43.1)	5696	
Median (IQR)	97 (85-110)		
Initial laboratory measures, median (IQR)[a]			
White blood cell count, $\times 10^9$/L	7.0 (5.2-9.5)	5680	3.8-10.5
Absolute count, $\times 10^9$/L			
Neutrophil	5.3 (3.7-7.7)	5645	1.8-7.4
Lymphocyte	0.88 (0.6-1.2)	5645	1.0-3.3
Lymphocyte, <1000 $\times 10^9$/L	3387 (60)		
Sodium, mmol/L	136 (133-138)	5645	135-145
Aspartate aminotransferase, U/L	46 (31-71)	5586	10-40
Aspartate aminotransferase >40 U/L	3263 (58.4)		
Alanine aminotransferase, U/L	33 (21-55)	5587	10-45
Alanine aminotransferase >60 U/L	2176 (39.0)		
Creatine kinase, U/L	171 (84-397)	2527	25-200
Venous lactate, mmol/L	1.5 (1.1-2.1)	2508	0.7-2.0
Troponin above test-specific upper limit of normal[b]	801 (22.6)	3533	
Brain-type natriuretic peptide, pg/mL	385.5 (106-1996.8)	1818	0-99
Procalcitonin, ng/mL	0.2 (0.1-0.6)	4138	0.02-0.10
D-dimer, ng/mL	438 (262-872)	3169	0-229
Ferritin, ng/mL	798 (411-1515)	4344	15-400
C-reactive protein, mg/dL	13.0 (6.4-26.9)	4517	0.0-0.40
Lactate dehydrogenase, U/L	404.0 (300-551.5)	4003	50-242
Admission studies[a]			
ECG, QTC >500[c]	260 (6.1)	4250	<400
Respiratory viral panel, positive for non-COVID-19 respiratory virus	42 (2.1)	1996	
Chlamydia pneumoniae	2 (4.8)		
Coronavirus (non-COVID-19)	7 (16.7)		
Entero/rhinovirus	22 (52.4)		
Human metapneumovirus	2 (4.8)		
Influenza A	1 (2.4)		
Mycoplasma pneumoniae	1 (2.4)		
Parainfluenza 3	3 (7.1)		
Respiratory syncytial virus	4 (9.5)		
Length of stay for patients in hospital at study end point, median (IQR), d	4.5 (2.4-8.1)		
No.	3066		

Abbreviations: COVID-19, coronavirus disease 2019; ECG, electrocardiogram; IQR, interquartile range; QTC, corrected QT interval.

SI conversion factors: To convert alanine aminotransferase, alkaline phosphatase, aspartate aminotransferase, creatinine kinase, and lactate dehydrogenase to μkat/L, multiply by 0.0167.

[a] Triage vital signs, initial laboratory measures, and admission studies were selected to be included here based on relevance to the characterization of patients with COVID-19.

[b] Troponin I; troponin T; and troponin T, high sensitivity are used at about equal frequency across these institutions. For simplicity, we present the number and percentage of test results that were above the upper limit of normal for the individual references ranges for these 3 tests.

[c] QTC resulted from the automated ECG reading.

co-infection with another respiratory virus for those tested was 2.1% (42/1996). Discharge disposition by 10-year age intervals of all 5700 study patients is included in **Table 4**. Length of stay for those who died, were discharged alive, and remained in hospital are presented as well. Among the 3066 patients who remained hospitalized at the final study follow-up date (median age, 65 years [IQR 54-75]), the median follow-up at time of censoring was 4.5 days (IQR, 2.4-8.1). Mortality was 0% (0/20) for male and female patients younger than 20 years. Mortality rates were higher for male compared with female patients at every 10-year age interval older than 20 years.

Table 3. Hospital Characteristics and Admission Rates

Hospital[a]	No. (%)		
	Study admissions (N = 5700)	Acute beds (March occupancy), mean[b]	Annual emergency department visits (% admitted)
North Shore University Hospital	1073 (18.8)	637 (92)	51 000 (34)
Long Island Jewish Medical Center	1151 (20.2)	517 (91)	66 000 (28)
Staten Island University Hospital	674 (11.9)	466 (85)	93 000 (25)
Lenox Hill Hospital	558 (9.8)	324 (75)	40 000 (29)
Southside Hospital	445 (7.8)	270 (86)	59 000 (18)
Huntington Hospital	359 (6.3)	231 (81)	40 000 (22)
Long Island Jewish Forest Hills	608 (10.7)	187 (86)	42 000 (21)
Long Island Jewish Valley Stream	355 (6.2)	180 (75)	31 000 (23)
Plainview Hospital	231 (4.1)	156 (70)	24 000 (29)
Cohen Children's Medical Center	42 (0.7)	111 (78)	48 000 (14)
Glen Cove Hospital, nonteaching	117 (2.1)	66 (78)	15 000 (20)
Syosset Hospital	87 (1.5)	55 (70)	12 000 (21)

[a] Teaching hospital unless otherwise noted.

[b] More than 1200 acute beds were added across the system during the month of March 2020.

Table 4. Discharge Disposition by 10-Year Age Intervals of Patients Hospitalized With COVID-19

	Patients discharged alive or dead at study end point						Patients in hospital at study end point	
	Died, No./No. (%)		Length of stay among those who died, median (IQR), d[a]	Discharged alive, No./No. (%)		Length of stay among those discharged alive, median (IQR), d[a]	No./No. (%)	Length of stay, median (IQR), d[a]
	Male	Female		Male	Female			
Age intervals, y								
0-9	0/13	0/13	NA	13/13 (100)	13/13 (100)	2.0 (1.7-2.7)	7/33 (21.2)	4.3 (3.1-12.5)
10-19	0/1	0/7	NA	1/1 (100)	7/7 (100)	1.8 (1.0-3.1)	9/17 (52.9)	3.3 (2.8-4.3)
20-29	3/42 (7.1)	1/55 (1.8)	4.0 (0.8-7.4)	39/42 (92.9)	54/55 (98.2)	2.5 (1.8-4.0)	52/149 (34.9)	3.2 (1.9-6.4)
30-39	6/130 (4.6)	2/81 (2.5)	2.8 (2.4-3.6)	124/130 (95.4)	79/81 (97.5)	3.7 (2.0-5.8)	142/353 (40.2)	5.1 (2.5-9.0)
40-49	19/233 (8.2)	3/119 (2.5)	5.6 (3.0-8.4)	214/233 (91.8)	116/119 (97.5)	3.9 (2.3-6.1)	319/671 (47.5)	4.9 (2.9-8.2)
50-59	40/327 (12.2)	13/188 (6.9)	5.9 (3.1-9.5)	287/327 (87.8)	175/188 (93.1)	3.8 (2.5-6.7)	594/1109 (53.6)	4.9 (2.8-8.0)
60-69	56/300 (18.7)	28/233 (12.0)	5.7 (2.6-8.2)	244/300 (81.3)	205/233 (88.0)	4.3 (2.5-6.8)	771/1304 (59.1)	5.0 (2.4-8.2)
70-79	91/254 (35.8)	54/197 (27.4)	5.0 (2.7-7.8)	163/254 (64.2)	143/197 (72.6)	4.6 (2.8-7.8)	697/1148 (60.7)	4.5 (2.3-8.2)
80-89	94/155 (60.6)	76/158 (48.1)	3.9 (2.1-6.5)	61/155 (39.4)	82/158 (51.9)	4.4 (2.7-7.7)	369/682 (54.1)	4.1 (2.1-7.4)
≥90	28/44 (63.6)	39/84 (46.4)	3.0 (0.7-5.5)	16/44 (36.4)	45/84 (53.6)	4.8 (2.8-8.4)	106/234 (45.3)	3.2 (1.5-6.4)

Abbreviations: COVID-19, coronavirus disease 2019; IQR, interquartile range; NA, not applicable.

[a] Length of stay begins with admission time and ends with discharge time, time at death, or midnight on the last day of data collection for the study. It does not include time in the emergency department.

Outcomes for Patients Who Were Discharged or Died

Among the 2634 patients who were discharged or had died at the study end point, during hospitalization, 373 (14.2%) were treated in the ICU, 320 (12.2%) received invasive mechanical ventilation, 81 (3.2%) were treated with kidney replacement therapy, and 553 (21%) died (**Table 5**). As of April 4, 2020, for patients requiring mechanical ventilation (n = 1151, 20.2%), 38 (3.3%) were discharged alive, 282 (24.5%) died, and 831 (72.2%) remained in hospital. Mortality rates for those who received mechanical ventilation in the 18-to-65 and older-than-65 age groups were 76.4% and 97.2%, respectively. Mortality rates for those in the 18-to-65 and older-than-65 age groups who did not receive mechanical ventilation were 1.98% and 26.6%, respectively. There were no deaths in the younger-than-18 age group. The overall length of stay was 4.1 days (IQR, 2.3-6.8). The median postdischarge follow-up time was 4.4 days (IQR, 2.2-9.3). A total of 45 patients (2.2%) were readmitted during the study period. The median time to readmission was 3 days (IQR, 1.0-4.5). Of the patients who were discharged or had died at the study end point, 436 (16.6%) were younger than age 50 with a score of 0 on the Charlson Comorbidity Index, of whom 9 died.

Outcomes by Age and Risk Factors

For both patients discharged alive and those who died, the percentage of patients who were treated in the ICU or received invasive mechanical ventilation was increased for the 18-to-65 age group compared with the older-than-65 years age group (Table 5). For patients discharged alive, the lowest absolute lymphocyte count during hospital course was lower for progressively older age groups. For patients discharged alive, the readmission rates and the percentage of patients discharged to a facility (such as a nursing home or

Table 5. Clinical Measures and Outcomes for Patients Discharged Alive, Dead, and In Hospital at Study End Point by Age

Clinical measure	Total discharged alive and dead patients (N = 2634)	Discharged alive			Died			In hospital		
		<18 y (n = 32)	18-65 y (n = 1373)	>65 y (n = 676)	<18 y (n = 0)	18-65 y (n = 134)	>65 y (n = 419)	<18 (n = 14)	18-65 (n = 1565)	>65 (n = 1487)
Invasive mechanical ventilation[a]	320 (12.2)	0	33 (2.4)	5 (0.7)	NA	107 (79.9)	175 (41.8)	4 (28.6)	449 (28.7)	378 (25.4)
ICU care	373 (14.2)	2 (6.3)	62 (4.5)	18 (2.7)	NA	109 (81.3)	182 (43.4)	5 (35.7)	490 (31.3)	413 (27.8)
Absolute lymphocyte count at nadir, median (IQR), $\times 10^9$/L (reference range, 1.0-3.3)	0.8 (0.5-1.14)	2.3 (1.2-5.0)	0.9 (0.7-1.2)	0.8 (0.5-1.1)	NA	0.5 (0.3-0.8)	0.5 (0.3-0.8)	2.0 (1.0-3.5)	0.7 (0.5-1.0)	0.6 (0.4-0.9)
No.	2626	32	1371	675		134	417	3	1564	1486
Acute kidney injury[b]	523 (22.2)	1 (11.1)	93 (7.5)	82 (13.1)	NA	98 (83.8)	249 (68.4)	2 (14.3)	388 (25.5)	457 (34.5)
No.	2351	8	1237	624		117	364	8	1400	1326
Kidney replacement therapy	81 (3.2)	0	2 (0.1)	1 (0.2)	NA	43 (35.0)	35 (8.8)	0	82 (5.4)	62 (4.4)
Acute hepatic injury[c]	56 (2.1)	0	3 (0.2)	0	NA	25 (18.7)	28 (6.7)	0	21 (1.3)	12 (0.8)
No.			1371	675		134	417	3	1564	1486
Outcomes										
Length of stay, median (IQR), d[d]	4.1 (2.3-6.8)	2.0 (1.7-2.8)	3.8 (2.3-6.2)	4.5 (2.7-7.2)	NA	5.5 (2.9-8.4)	4.4 (2.1-7.1)	4.0 (2.4-6.2)	4.8 (2.5-8.1)	4.4 (2.3-8.0)
Discharged alive	3.9 (2.4-6.7)									
Died	4.8 (2.3-7.4)									
Died	553 (21)	NA	NA	NA	NA	NA	NA	NA	NA	N/A
Died, of those who did not receive mechanical ventilation	271/2314 (11.7)	NA	NA	NA	NA	NA	NA	NA	NA	
Died, of those who did receive mechanical ventilation	282/320 (88.1)									
Readmitted[e]	45 (2.2)	1 (3.1)	22 (1.6)	22 (3.3)	NA	NA	NA	NA	NA	NA
Discharge disposition of 2081 patients discharged alive										
No.	2081									
Home	1959 (94.1)	32 (100)	1345 (98.0)	582 (86.1)	NA	NA	NA	NA	NA	NA
Facilities (ie, nursing home, rehab)	122 (5.9)	0	28 (2.0)	94 (13.9)	NA	NA	NA	NA	NA	NA

Abbreviations: ICU, intensive care unit; IQR, interquartile range; NA, not applicable.

[a] Policy in the system has been not to treat patients with COVID-19 with bilevel positive airway pressure and continuous positive airway pressure out of concern for aerosolizing virus particles and therefore that information is not reported here.

[b] Acute kidney injury was identified as an increase in serum creatinine by ≥0.3 mg/dL (≥26.5 mol/L) within 48 hours or an increase in serum creatinine to ≥1.5 times baseline within the prior 7 days compared with the preceding 1 year of data in acute care medical records. Acute kidney injury is calculated only for patients with record of baseline kidney function data available and without a diagnosis of end-stage kidney disease.

[c] Acute hepatic injury was defined as an elevation in aspartate aminotransferase or alanine aminotransferase of >15 times the upper limit of normal.

[d] Length of stay begins with admission time and ends with discharge time or time of death. It does not include time in the emergency department.

[e] Data are presented here for readmission during the study period, March 1 to April 4, 2020.

rehabilitation), as opposed to home, increased for progressively older age groups.

Of the patients who died, those with diabetes were more likely to have received invasive mechanical ventilation or care in the ICU compared with those who did not have diabetes (eTable 1 in the Supplement). Of the patients who died, those with hypertension were less likely to have received invasive mechanical ventilation or care in the ICU compared with those without hypertension. The percentage of patients who developed acute kidney injury was increased in the subgroups with diabetes compared with subgroups without those conditions.

Angiotensin-Converting Enzyme Inhibitor and Angiotensin II Receptor Blocker Use

Home medication reconciliation information was available for 2411 (92%) of the 2634 patients who were discharged or who died by the study end. Of these 2411 patients, 189 (7.8%) were taking

an angiotensin-converting enzyme inhibitor (ACEi) at home and 267 (11.1%) were taking an angiotensin II receptor blocker (ARB) at home. The median number of total home medications was 3 (IQR, 0-7). Outcomes for subgroups of patients with hypertension by use of ACEi or ARB home medication are shown in eTable 2 in the Supplement. Numbers provided for total patients taking ACEi or ARB therapy in eTable 2 in the Supplement are provided only for patients who also had a diagnosis of hypertension.

Of the patients taking an ACEi at home, 91 (48.1%) continued taking an ACEi while in the hospital and the remainder discontinued this type of medication during their hospital visit. Of the patients taking an ARB at home, 136 (50.1%) continued taking an ARB while in the hospital and the remainder discontinued taking this type of medication during their hospital visit. Of patients who were not prescribed an ACEi or ARB at home, 49 started treatment with an ACEi and 58 started treatment with an ARB during their hospitalization. Mortality rates for patients with hypertension not taking an ACEi or ARB, taking an ACEi, and taking an ARB were 26.7%, 32.7%, and 30.6%, respectively.

Discussion

To our knowledge, this study represents the first large case series of sequentially hospitalized patients with confirmed COVID-19 in the US. Older persons, men, and those with preexisting hypertension and/or diabetes were highly prevalent in this case series and the pattern was similar to data reported from China.[4] However, mortality rates in this case series were significantly lower, possibly due to differences in thresholds for hospitalization. This study reported mortality rates only for patients with definite outcomes (discharge or death), and longer-term study may find different mortality rates as different segments of the population are infected. The findings of high mortality rates among ventilated patients are similar to smaller case series reports of critically ill patients in the US.[10]

ACEi and ARB medications can significantly increase mRNA expression of cardiac angiotensin-converting enzyme 2 (ACE2),[11] leading to speculation about the possible adverse, protective, or biphasic effects of treatment with these medications.[12] This is an important concern because these medications are the most prevalent antihypertensive medications among all drug classes.[13] However, this case series design cannot address the complexity of this question, and the results are unadjusted for known confounders, including age, sex, race, ethnicity, socioeconomic status indicators, and comorbidities such as diabetes, chronic kidney disease, and heart failure.

Mortality rates are calculated only for patients who were discharged alive or died by the study end point. This biases our rates toward including more patients who died early in their hospital course. Most patients in this study were still in hospital at the study end point (3066, 53.8%). We expect that as these patients complete their hospital course, reported mortality rates will decline.

Limitations

This study has several limitations. First, the study population only included patients within the New York metropolitan area. Second, the data were collected from the electronic health record database. This precluded the level of detail possible with a manual medical record review. Third, the median postdischarge follow-up time was relatively brief at 4.4 days (IQR, 2.2-9.3). Fourth, subgroup descriptive statistics were unadjusted for potential confounders. Fifth, clinical outcome data were available for only 46.2% of admitted patients. The absence of data on patients who remained hospitalized at the final study date may have biased the findings, including the high mortality rate of patients who received mechanical ventilation older than age 65 years.

Conclusions

This case series provides characteristics and early outcomes of sequentially hospitalized patients with confirmed COVID-19 in the New York City area.

ARTICLE INFORMATION

Accepted for Publication: April 16, 2020.

Published Online: April 22, 2020. doi:10.1001/jama.2020.6775

Correction: This article was corrected on April 24, 2020, to clarify the mortality rate of ventilated patients, correct the COVID-19 positive/negative test results, and correct the data for concurrent entero/rhinovirus infection in Table 2.

The Northwell COVID-19 Research Consortium Authors: Douglas P. Barnaby, MD, MSc; Lance B. Becker, MD; John D. Chelico, MD, MA; Stuart L. Cohen, MD; Jennifer Cookingham, MHA; Kevin Coppa, BS; Michael A. Diefenbach, PhD; Andrew J. Dominello, BA; Joan Duer-Hefele, RN, MA; Louise Falzon, BA, PGDipInf; Jordan Gitlin, MD; Negin Hajizadeh, MD, MPH; Tiffany G. Harvin, MBA; David A. Hirschwerk, MD; Eun Ji Kim, MD, MS, MS; Zachary M. Kozel, MD; Lyndonna M. Marrast, MD, MPH; Jazmin N. Mogavero, MA; Gabrielle A. Osorio, MPH; Michael Qiu, MD, PhD; Theodoros P. Zanos, PhD.

Affiliations of The Northwell COVID-19 Research Consortium Authors: Institute of Health Innovations and Outcomes Research, Feinstein Institutes for Medical Research, Northwell Health, Manhasset, New York (Barnaby, Chelico, Cohen, Cookingham, Diefenbach, Dominello, Duer-Hefele, Falzon, Hajizadeh, Harvin, Kim, Marrast, Mogavero, Osorio); Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Northwell Health, Hempstead, New York (Barnaby, Becker, Chelico, Cohen, Gitlin, Hajizadeh, Hirschwerk, Kim, Kozel, Marrast); Department of Information Services, Northwell Health, New Hyde Park, New York (Coppa, Qiu); Institute of Bioelectronic Medicine, Feinstein Institutes for Medical Research, Northwell Health, Manhasset, New York (Zanos).

Author Contributions: Drs Richardson and Davidson had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Richardson, Hirsch, Narasimhan, Crawford, McGinn, Davidson, Barnaby, Chelico, Cohen, Cookingham, Coppa, Diefenbach, Duer-Hefele, Dominello, Falzon, Gitlin, Hirschwerk, Kozel, Marrast, Mogavero.
Acquisition, analysis, or interpretation of data: Richardson, Hirsch, Narasimhan, Crawford, Davidson, Barnaby, Becker, Chelico, Cohen, Coppa, Diefenbach, Duer-Hefele, Hajizadeh, Harvin, Hirschwerk, Kim, Kozel, Marrast, Osorio, Qiu, Zanos.
Drafting of the manuscript: Richardson, McGinn, Davidson, Cookingham, Falzon, Harvin, Mogavero, Qiu.
Critical revision of the manuscript for important intellectual content: Richardson, Hirsch, Narasimhan, Crawford, McGinn, Barnaby, Becker, Chelico, Cohen, Coppa, Diefenbach, Duer-Hefele, Dominello, Gitlin, Hajizadeh, Hirschwerk, Kim, Kozel, Marrast, Osorio, Zanos.
Statistical analysis: Hirsch, Chelico, Zanos.
Obtained funding: Richardson.
Administrative, technical, or material support: Richardson, Narasimhan, Crawford, Davidson, Chelico, Cookingham, Diefenbach, Dominello, Harvin, Mogavero, Osorio, Zanos.

Supervision: Narasimhan, McGinn, Becker, Chelico, Zanos.

Conflict of Interest Disclosures: Dr Crawford reported receiving grants from Regeneron outside the submitted work. Dr Becker reported serving on the scientific advisory board for Nihon Kohden and receiving grants from the National Institutes of Health, United Therapeutics, Philips, Zoll, and Patient-Centered Outcomes Research Institute outside the submitted work. Dr Cohen reported receiving personal fees from Infervision outside the submitted work. No other disclosures were reported.

Funding/Support: This work was supported by grants R24AG064191 from the National Institute on Aging of the National Institutes of Health; R01LM012836 from the National Library of Medicine of the National Institutes of Health; and K23HL145114 from the National Heart, Lung, and Blood Institute.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

The Northwell COVID-19 Research Consortium Investigators: Douglas P. Barnaby, MD, MSc, Lance B. Becker, MD, John D. Chelico, MD, MA, Stuart L. Cohen, MD, Jennifer Cookingham, MHA, Kevin Coppa, BS, Michael A. Diefenbach, PhD, Andrew J. Dominello, BA, Joan Duer-Hefele, RN, MA, Louise Falzon, BA, Jordan Gitlin, MD, Negin Hajizadeh, MD, MPH, Tiffany G. Harvin, MBA, David A. Hirschwerk, MD, Eun Ji Kim, MD, MS, MS, Zachary M. Kozel, MD, Lyndonna M. Marrast, MD, MPH, Jazmin N. Mogavero, MA, Gabrielle A. Osorio, MPH, Michael Qiu, MD, PhD, and Theodoros P. Zanos, PhD.

Disclaimer: The views expressed in this article are those of the authors and do not represent the views of the National Institutes of Health, the US Department of Health and Human Services, or any other government entity. Karina W. Davidson is a member of the US Preventive Services Task Force (USPSTF). This article does not represent the views and policies of the USPSTF.

REFERENCES

1. Holshue ML, DeBolt C, Lindquist S, et al; Washington State 2019-nCoV Case Investigation Team. First case of 2019 novel coronavirus in the United States. *N Engl J Med*. 2020;382(10):929-936. doi:10.1056/NEJMoa2001191

2. The Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Coronavirus COVID-19 global cases. Accessed March 30, 2020. https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6

3. Centers for Disease Control and Prevention. Coronavirus disease 2019 (COVID-19): cases in US. Accessed March 25, 2020. https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html

4. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395(10229):1054-1062. doi:10.1016/S0140-6736(20)30566-3

5. United Nations Department of Economic and Social Affairs Population Dynamics. World population prospects 2019. Accessed April 6, 2020. https://population.un.org/wpp/Graphs/DemographicProfiles/Pyramid/840

6. Chen Z, Peto R, Zhou M, et al; China Kadoorie Biobank (CKB) collaborative group. Contrasting male and female trends in tobacco-attributed mortality in China: evidence from successive nationwide prospective cohort studies. *Lancet*. 2015;386(10002):1447-1456. doi:10.1016/S0140-6736(15)00340-2

7. Institute for Health Metrics and Evaluation. GBD Compare/Viz Hub. Accessed April 6, 2020. https://vizhub.healthdata.org/gbd-compare/

8. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373-383. doi:10.1016/0021-9681(87)90171-8

9. Kellum JA, Lameire N, Aspelin P, et al. Kidney Disease: Improving Global Outcomes (KDIGO) acute kidney injury work group: KDIGO clinical practice guideline for acute kidney injury. *Kidney Int Suppl*. 2012;2(1):1-138.

10. Arentz M, Yim E, Klaff L, et al. Characteristics and outcomes of 21 critically ill patients with COVID-19 in Washington State. *JAMA*. 2020. doi:10.1001/jama.2020.4326

11. Ferrario CM, Jessup J, Chappell MC, et al. Effect of angiotensin-converting enzyme inhibition and angiotensin II receptor blockers on cardiac angiotensin-converting enzyme 2. *Circulation*. 2005;111(20):2605-2610. doi:10.1161/CIRCULATIONAHA.104.510461

12. Sommerstein R, Kochen MM, Messerli FH, Gräni C. Coronavirus disease 2019 (COVID-19): do angiotensin-converting enzyme inhibitors/angiotensin receptor blockers have a biphasic effect? *J Am Heart Assoc*. 2020;9(7):e016509. doi:10.1161/JAHA.120.016509

13. Derington CG, King JB, Herrick JS, et al. Trends in antihypertensive medication monotherapy and combination use among US adults, National Health and Nutrition Examination Survey 2005–2016. *Hypertension*. 2020;75(4):973-981. doi:10.1161/HYPERTENSIONAHA.119.14360

JAMA | Original Investigation

The State of US Health, 1990-2016
Burden of Diseases, Injuries, and Risk Factors Among US States

The US Burden of Disease Collaborators

INTRODUCTION Several studies have measured health outcomes in the United States, but none have provided a comprehensive assessment of patterns of health by state.

OBJECTIVE To use the results of the Global Burden of Disease Study (GBD) to report trends in the burden of diseases, injuries, and risk factors at the state level from 1990 to 2016.

DESIGN AND SETTING A systematic analysis of published studies and available data sources estimates the burden of disease by age, sex, geography, and year.

MAIN OUTCOMES AND MEASURES Prevalence, incidence, mortality, life expectancy, healthy life expectancy (HALE), years of life lost (YLLs) due to premature mortality, years lived with disability (YLDs), and disability-adjusted life-years (DALYs) for 333 causes and 84 risk factors with 95% uncertainty intervals (UIs) were computed.

RESULTS Between 1990 and 2016, overall death rates in the United States declined from 745.2 (95% UI, 740.6 to 749.8) per 100 000 persons to 578.0 (95% UI, 569.4 to 587.1) per 100 000 persons. The probability of death among adults aged 20 to 55 years declined in 31 states and Washington, DC from 1990 to 2016. In 2016, Hawaii had the highest life expectancy at birth (81.3 years) and Mississippi had the lowest (74.7 years), a 6.6-year difference. Minnesota had the highest HALE at birth (70.3 years), and West Virginia had the lowest (63.8 years), a 6.5-year difference. The leading causes of DALYs in the United States for 1990 and 2016 were ischemic heart disease and lung cancer, while the third leading cause in 1990 was low back pain, and the third leading cause in 2016 was chronic obstructive pulmonary disease. Opioid use disorders moved from the 11th leading cause of DALYs in 1990 to the 7th leading cause in 2016, representing a 74.5% (95% UI, 42.8% to 93.9%) change. In 2016, each of the following 6 risks individually accounted for more than 5% of risk-attributable DALYs: tobacco consumption, high body mass index (BMI), poor diet, alcohol and drug use, high fasting plasma glucose, and high blood pressure. Across all US states, the top risk factors in terms of attributable DALYs were due to 1 of the 3 following causes: tobacco consumption (32 states), high BMI (10 states), or alcohol and drug use (8 states).

CONCLUSIONS AND RELEVANCE There are wide differences in the burden of disease at the state level. Specific diseases and risk factors, such as drug use disorders, high BMI, poor diet, high fasting plasma glucose level, and alcohol use disorders are increasing and warrant increased attention. These data can be used to inform national health priorities for research, clinical care, and policy.

Editorial page 1438

Author Audio Interview

Supplemental content

CME Quiz at jamanetwork.com/learning and CME Questions page 1503

Group Information: The US Burden of Disease Collaborators are listed at the end of this article.

Corresponding Author: Christopher J. L. Murray, MD, DPhil, Institute for Health Metrics and Evaluation, University of Washington, 2301 5th Ave, Ste 600, Seattle, WA 98121 (cjlm@uw.edu).

JAMA. 2018;319(14):1444-1472. doi:10.1001/jama.2018.0158

Previous studies have reported on health disparities in US states and counties.[1,2] These studies showed that health disparities have increased with time. Recent attention has focused on increased mortality in some age groups and a decline in life expectancy.[3] In addition, the performance of the US health care system does not match its level of spending on health and lags behind countries with similar financial resources.[4] For example, in 2014, US life expectancy ranked 43rd in the world, although the United States spent the most ($3.0 trillion) on health care, exceeding the median amount spent by Organisation for Economic Co-operation and Development countries by 35%.[5]

Several studies have shown large variations in risk factors by state and county, and these variations have contributed to differences in health outcomes.[6-9] In the Global Burden of Disease Study 2010 (GBD 2010) US Burden of Disease report, the following risk factors were reported as the main causes associated with US morbidity and mortality (percent contributed to total disability-adjusted life-years [DALYs] in 2010): poor diet (14%), smoking (11%), high blood pressure (8%), and obesity (11%).[10-12] None of the previous studies of US health have been as comprehensive as the GBD study.[5,13-18] The GBD systematically accounts for differences in data sources and biases and analyzes levels and trends for causes and risk factors within the same computational framework, which maximizes comparability across states, years, and different age groups by sex. GBD is now conducted on an annual cycle, with GBD 2016 providing updated estimates of mortality, morbidity, and risk factors in 195 locations, including the United States, from 1990 to 2016.

The findings of GBD 2016 indicate that while the United States overall is experiencing improvements in health outcomes, the patterns of health burden at the state level vary across geography. Routinely monitoring the trend of burden of disease at the state level is essential given the vital role of states in many aspects of health and social policy[19]—from the Medicaid program to regulation of private insurers[20] and considering that individual states also experience different economic circumstances. The current study uses GBD 2016 to report the change in burden of disease, including injuries and risk factors at the state level, from 1990 to 2016.

Key Points

Question How have the levels and trends of burden of diseases, injuries, and risk factors in the United States changed from 1990 to 2016 by state?

Findings This study, involving examination of 333 causes and 84 risk factors, demonstrated that health in the United States improved from 1990 to 2016, although the drivers of mortality and morbidity have changed in some states, with specific risk factors such as drug use disorders, high body mass index (BMI), and alcohol use disorders being associated with adverse outcomes. In 5 states, the probability of death between ages 20 and 55 years has increased more than 10% between 1990 and 2016.

Meaning Differences in health outcomes and drivers of morbidity and mortality at the state level indicate the need for greater investment in preventive and medical care across the life course. The intersection of risk, mortality, and morbidity in particular geographic areas needs to be further explored at the state level.

Methods

Overview

The GBD study is estimated annually and each round of results is internally consistent (cause-specific mortality estimates are summed to match all-cause mortality estimates) and collectively exhaustive (residual categories ["other"] are captured to enable quantifying total burden) (Sections 1-5 in Supplement 1). The numbers reported in the previous round of GBD are not identical to those of the current round (GBD 2016) for 2 main reasons. First, since the GBD 2010 Special Communication regarding US risk factors, there has been further refinement of the "garbage coding" (ie, ill-defined causes of death) redistribution methods (Supplement 1). Second, the new analysis at the state level changes some of the estimation slightly when aggregated to the national level. GBD 2016 provides a new time series.

The GBD 2016 methodology has been published previously.[5,13-18] GBD uses several metrics to report results for health loss related to specific diseases, injuries, and risk factors: deaths and death rates, years of life lost (YLLs) due to premature mortality, prevalence and prevalence rates for sequelae, years lived with disability (YLDs), and DALYs (**Box**; Sections 2, 3, and 5 in Supplement 1 and Appendix Table 2 in Supplement 2). GBD provides a comprehensive assessment of all-cause mortality and estimates for death due to 264 causes in 195 countries and territories from 1990 to 2016, as well as 333 causes of DALYs (Appendix Table 2 in Supplement 2). GBD 2016 has 4 levels of causes that are mutually exclusive (Appendix Table 3 in Supplement 2). Level 1 has 3 causes: communicable, maternal, neonatal, and nutritional disorders; noncommunicable diseases; and injuries. Level 2 has 21 causes. Levels 3 and 4 consist of more disaggregated causes. GBD 2016 documented each step of the estimation processes, as well as data sources, in accordance with the Guidelines for Accurate and Transparent Health Estimates Reporting.[21] A more detailed methodology is available in the appendix to this article (Sections 8 and 9 in Supplement 1).

Data

To estimate the US burden of disease prevalence, computation for each sequela began with a systematic analysis of published studies and available data sources providing information on prevalence, incidence, remission, and excess mortality, such as the National Health and Nutrition Examination Surveys,[22] state inpatient databases,[23] the National Ambulatory Medical Care Survey,[24] National Hospital Ambulatory Medical Care Survey,[25] Medical Expenditure Panel Survey,[26] National Comorbidity Survey,[27] National Epidemiological Survey on Alcohol and Related Conditions,[28] National Survey on Drug Use and Health,[29] US Department of Agriculture Continuing Survey of Food Intakes,[30] Marketscan,[31] National Health Interview Survey,[32] Behavioral Risk Factor Surveillance System,[33] and the Centers for Disease Control and Prevention Disease Surveillance Reports.[34]

Hospital inpatient data were extracted and used for this analysis. Moreover, outpatient encounter data were available for the United States through aggregate data derived from a

Box. Glossary of Terms

Disability-adjusted life-years: a summary metric of population health. DALYs represent a health gap and, as such, measure the state of a population's health compared to a normative goal. The goal is for individuals to live the standard life expectancy in full health. DALYs are the sum of 2 components: years of life lost (YLLs) and years lived with disability (YLDs).

Healthy life expectancy: the number of years that a person at a given age can expect to live in good health, taking into account mortality and disability.

Summary Exposure Value: the relative risk–weighted prevalence of exposure (developed for Global Burden of Diseases Study 2015).

Years lived with disability: computed as the prevalence of different disease sequelae and injury sequelae multiplied by disability weights for that sequela. Disability weights are selected on the basis of surveys of the general population about the health loss associated with the health state related to the disease sequela.

Years of life lost due to premature mortality: computed by multiplying the number of deaths at each age by a standard life expectancy at that age. The standard selected represents the normative goal for survival and has been computed based on the lowest recorded death rates across countries in 2010.

database of claims information for US private and public insurance schemes for the years 2000, 2010, and 2012. GBD methodology applied several correction factors to account for bias in health service encounter data from these claims that were available as aggregated by *International Classification of Diseases* (*ICD*) code and by primary diagnosis only. First, for chronic disorders, the study estimated the ratio between prevalence from primary diagnoses and prevalence from all diagnoses associated with a claim. Second, the claims data were used to generate the mean number of outpatient visits per disorder. Similarly, the study generated per-person discharge rates from hospital inpatient data in the United States.

All-Cause Mortality and Cause of Death

All-cause mortality was estimated by age, sex, geography, and year using 6 modeling approaches to assess cause-specific mortality; the Cause of Death Ensemble Model was used to generate estimates for the vast majority of causes. This analysis used deidentified death records from the National Center for Health Statistics (NCHS)[35] and population counts from the US Census Bureau,[36] NCHS, and the Human Mortality Database.[37] Deaths and population were tabulated by county, age group, sex, year, and (in the case of death data) cause. The cause list developed for the GBD[13] is arranged hierarchically in 4 levels. Within each level, the cause list is designed such that all deaths are assigned exactly 1 cause. As part of the GBD study, a map has been developed that allows *ICD-9* and *ICD-10* codes to be translated to GBD causes.

Previous studies have documented the existence of insufficiently specific or implausible causes of death used in death registration data that may lead to misleading geographic and temporal patterns.[38] Algorithms developed for the GBD were used to reallocate deaths assigned one of these garbage codes to plausible alternatives.[39] First, plausible target causes were assigned to each garbage code or group of garbage codes. Second, deaths were reassigned to specified target codes according to proportions derived in 1 of 4 ways: (1) published literature or expert opinion; (2) regression models; (3) according to the proportions initially observed among targets; and (4) for HIV/AIDS specifically, by comparison to years before HIV/AIDS became widespread. More detail on each of these methods is provided in Section 2 of Supplement 1.

Based on standard GBD methods, YLLs were computed by multiplying the number of deaths from each cause in each age group by the reference life expectancy at the mean of age of death among those who died in the age group. The YLLs computation is based on the precedent set by GBD and uses the same life table standard for calculating YLLs in all locations and years (essential for comparing estimates of YLLs across locations and years). The standard is meant to represent the mortality experience of a population with minimal excess mortality using the lowest observed age-specific mortality rates in 2016 among all countries with a population greater than 5 million. This standard does not vary with time because for most populations, the number of YLLs (once normalized for population size) is larger in earlier years than in later years due to improving survival rather than an artifact of the standard used.

Analysis of Incidence, Prevalence, and YLDs

In this study, incidence and prevalence of diseases by age, sex, cause, year, and geography were estimated using a wide range of updated and standardized analytical procedures. GBD used DisMod-MR, a Bayesian meta-regression tool, to determine prevalence and incidence by cause and sequelae.[40]

Data sources used for quantifying nonfatal outcomes are available online in the GBD results tool[41] and in Section 3 of Supplement 1. Prevalence of each sequela was multiplied by the disability weight for the corresponding health state to calculate YLDs for the sequela. The sum of all YLDs for relevant sequelae equated to overall YLDs for each disease. Details on disability weights for GBD 2016, including data collection and disability weight construction, are described elsewhere.[14]

Analysis of DALYs and HALE

Following GBD 2016 methods, national and state-level DALYs were computed by summing YLLs and YLDs for each cause, age, and sex in 1990, 1995, 2000, 2005, 2010, and 2016 (Section 4 in Supplement 1). DALYs were computed for 333 causes, with 95% uncertainty intervals (UIs) capturing the uncertainty for both YLL and YLD rates. HALE was calculated using the Sullivan method and generated 95% UIs that indicated uncertainty for age-specific death rates and YLDs per capita for each geography, age group, sex, and year. HALE was calculated for the United States and for each individual state using multiple-decrement life tables and estimated YLDs per capita; additional details on HALE methodology are provided in Section 4 in Supplement 1.

Risk Factors

GBD 2016 used the comparative risk assessment framework to estimate attributable deaths, DALYs, and trends in exposure by age group, sex, year, and geography for risks from 1990 to 2016. GBD has 84 behavioral, environmental and occupational, and metabolic risks or clusters of risks (Section 5 in Supplement 1). Risk-outcome pairs were included in the GBD 2016 study if they

met World Cancer Research Fund criteria for convincing or probable evidence. Relative risk (RR) estimates were extracted from published and unpublished randomized clinical trials, cohorts, and pooled cohorts. Risk exposures were estimated based on published studies, household surveys, US Census data, satellite data, and other sources. Two modeling approaches, a Bayesian meta-regression model and a spatiotemporal Gaussian process regression model, were developed for the GBD study and used to pool data from different sources, adjust for bias in the data, and incorporate potential covariates. GBD used the counterfactual scenario of theoretical minimum risk exposure level (ie, the level for a given risk exposure that could minimize risk at the population level) to attribute burden. A summary exposure value was developed for GBD 2015 as the RR-weighted prevalence of exposure (range, 0 [no excess risk exists in a population] to 1 [population is at the highest risk]).[16]

$$SEV = \frac{\sum_{i=1}^{n} Pr_i RR_i - 1}{RR_{max} - 1}$$

Where Pr_i is prevalence of category i exposure; RR_i is the RR of the category i; and RR_{max} is the maximum RR observed (between categories). This quantity is estimated for each age group, sex, geography, and year. In the case of dichotomous exposure, summary exposure value is equal to prevalence. For continuous risks, summary exposure value is defined as follows:

$$SEV = \frac{\int_{x=1}^{u} RR(x)\, P(x)\, dx - 1}{RR_{max} - 1}$$

Where (χ) is the density of exposure at level χ of exposure; (χ) is the RR of the level χ; and RR_{max} is the highest RR that is supported by data and reflects a level in which more than 1% of the population is exposed globally. In this study, summary exposure value is reported on a scale from 0% to 100% to emphasize that it is risk-weighted prevalence.

To calculate risk-attributable fractions of disease burden by cause, the effects of risk exposure levels were modeled, RRs associated with risk exposure and specific health outcomes were documented, and counterfactual levels of risk exposure on estimates of national and state-level deaths, YLLs, YLDs, and DALYs were computed. Detailed descriptions of the GBD 2016 methods for risk factor assessment and attribution are published elsewhere (Section 5 in Supplement 1).[5,13-18]

Decomposition of Changes in Probability of Death

The probability of death was calculated for 3 summary age intervals and the cause-specific contributions to each of these summary indicators for ages 0 to 20, 20 to 55, and 55 to 90 years. These age groups were chosen to reflect variations in trends and burden for adolescents, young adults, and older people. For each probability of death, the multiple decrement life-table method was used to compute the probability of death from each cause and the overall contribution of each cause of death to the summary probability of death. Although discrete age categories from life table calculations were used, the age categories slightly overlap for calculations of probability of death (ages 20 years and 55 years; see Section 6 in Supplement 1). To decompose the key drivers of life loss, the probability of death was determined and examined in parallel to the cause fractions for that same age group. Additional information on the decomposition of changes in the probability of death, including the formulas used, is available in the online methods section (Supplement 1).

Sociodemographic Index

GBD 2015 created a summary indicator that combines measures of income per capita, educational attainment for age 15 years or older, and total fertility rates.[39,42-46] This indicator is updated for each GBD round. The current sociodemographic index (SDI) was used to compare observed patterns of health loss to expected patterns for countries or locations with similar SDI scores. The SDI was computed similarly to the computation of the human development index to improve interpretability. Each component of the SDI was weighted equally and rescaled (range, 0 [lowest observed value during 1980-2016] to 1 [highest observed value during 1980-2016]). In the United States in 2016, the SDI ranged from 0.874 in Mississippi to 0.978 in Washington, DC (global SDI values in 2016 ranged from 0.268 in Somalia to 0.978 in Washington, DC).

Results

US Mortality and YLLs

Table 1 lists the 25 leading causes of death and premature mortality from 1990 to 2016. Ischemic heart disease (IHD); cancer of the trachea, bronchus, and lung; chronic obstructive pulmonary disease; Alzheimer disease and other dementias; and cancer of the colon and rectum were the 5 leading causes of death. Despite a 50.7% decline in age-standardized mortality and a 50.4% decline in age-standardized YLLs, IHD remained the leading cause of death and premature mortality. There was an increase in age-standardized mortality and in age-standardized YLLs from 1990 for chronic obstructive pulmonary disease (13.8% for deaths and 4.6% for YLLs) and for Alzheimer disease and other dementias (11.6% for deaths and 5.5% for YLLs). There was a decrease in age-standardized mortality and in age-standardized YLLs for colon and rectal cancer (29.6% for deaths and 27.9% for YLLs) and for breast cancer (32.6% for deaths and 36.0% for YLLs). Deaths from endocrine, metabolic, blood, and immune disorders increased by 89.1%, and YLLs increased by 60.3% from 1990 to 2016 (an increase in rank from 37 in 2010 to 22 in 2016). Other notable findings seen in Table 1 are declines in deaths from self-harm by firearm (13.2%) and physical violence by firearm (28.5%) but an increase in self-harm by other means (16.9%).

US YLDs

Table 2 provides the 25 leading diseases and injuries contributing to YLDs. Low back pain and major depressive disorders remained the first and second causes of YLDs in 2016. Age-standardized rates of low back pain declined by 12.4%, while age-standardized rates of major depressive disorder did not change from 1990. Diabetes, which was the third leading cause of YLDs, had a 29.6% increase in age-standardized rates from 1990, and increased in rank from 8 in 1990 to 3 in 2016. From 1990 to 2016,

Table 1. Deaths and Years of Life Lost With Percentage Change From the 25 Leading Causes of Death in the United States, 1990 and 2016

Diseases and Injuries	YLL Rank		YLLs, No. in Thousands (95% UI)		Percent Change		Deaths, No. in Thousands (95% UI)		Percent Change	
	1990	2016	1990	2016	YLLs	Age-Standardized YLL Rate	1990	2016	Deaths	Age-Standardized Death Rate
Ischemic heart disease	1	1	9445.4 (9309.4 to 9657.0)	7605.3 (7409.6 to 7802.4)	−19.5 (−21.7 to −17.3)	−50.4 (−51.8 to −49.1)	640.9 (632.7 to 653.3)	544.8 (531.7 to 557.5)	−15.0 (−17.1 to −13.0)	−50.7 (−52.0 to −49.5)
Tracheal, bronchus, and lung cancer	2	2	3155.4 (3102.5 to 3210.7)	3586.1 (3493.4 to 3681.9)	13.6 (10.3 to 17.2)	−32.8 (−34.8 to −30.6)	151.0 (148.5 to 153.6)	191.5 (186.5 to 196.8)	26.8 (22.9 to 31.0)	−24.0 (−26.3 to −21.6)
Chronic obstructive pulmonary disease	4	3	1382.5 (1326.7 to 1412.8)	2347.4 (2267.8 to 2463.5)	69.8 (63.2 to 79.7)	4.6 (0.6 to 10.6)	86.9 (83.6 to 88.8)	163.8 (158.2 to 172.1)	88.5 (80.9 to 100.0)	13.8 (9.2 to 20.7)
Alzheimer disease and other dementias	7	4	1049.8 (913.6 to 1224.1)	1875.8 (1690.2 to 2076.6)	78.7 (66.5 to 91.8)	5.5 (−1.0 to 12.7)	116.4 (100.7 to 135.1)	238.9 (214.3 to 264.8)	105.3 (90.6 to 119.9)	11.6 (4.4 to 19.1)
Colon and rectum cancer	6	5	1241.6 (1219.5 to 1264.9)	1437.0 (1393.5 to 1482.4)	15.7 (11.5 to 20.0)	−27.9 (−30.5 to −25.2)	68.4 (67.2 to 69.7)	79.3 (77.0 to 81.8)	15.9 (11.7 to 20.1)	−29.6 (−32.1 to −27.0)
Motor vehicle road injuries	3	6	1792.7 (1696.8 to 1953.0)	1349.9 (1275.6 to 1484.0)	−24.7 (−27.8 to −20.8)	−39.4 (−41.9 to −36.2)	36.1 (34.0 to 39.1)	31.3 (29.5 to 34.4)	−13.4 (−16.7 to −9.0)	−35.4 (−37.8 to −32.0)
Lower respiratory infections	8	7	1044.8 (1006.1 to 1081.6)	1334.8 (1268.2 to 1398.1)	27.8 (21.8 to 33.8)	−19.0 (−22.6 to −15.3)	63.5 (59.9 to 66.9)	96.0 (89.6 to 102.3)	51.2 (43.4 to 59.6)	−12.3 (−16.5 to −7.8)
Diabetes	12	8	921.1 (901.8 to 940.9)	1305.7 (1264.1 to 1345.5)	41.8 (36.5 to 47.0)	−10.1 (−13.4 to −6.8)	49.2 (48.1 to 50.3)	71.5 (69.2 to 73.7)	45.3 (40.0 to 50.4)	−11.4 (−14.6 to −8.2)
Intracerebral hemorrhage	13	9	901.0 (872.3 to 938.0)	1152.5 (1109.5 to 1199.2)	27.9 (22.3 to 33.1)	−19.3 (−22.7 to −16.0)	42.2 (40.5 to 44.1)	63.9 (61.1 to 67.1)	51.3 (44.8 to 58.0)	−10.3 (−13.9 to −6.4)
Ischemic stroke	11	10	1006.2 (959.6 to 1050.4)	1139.8 (1085.3 to 1194.3)	13.3 (8.7 to 17.6)	−31.2 (−34.0 to −28.6)	84.0 (79.1 to 88.4)	113.3 (106.4 to 120.4)	34.8 (28.8 to 40.7)	−23.8 (−27.0 to −20.9)
Breast cancer	10	11	1023.4 (996.5 to 1051.8)	1056.8 (1014.5 to 1104.5)	3.3 (−1.5 to 8.9)	−36.0 (−39.0 to −32.5)	43.9 (42.7 to 45.1)	49.3 (47.4 to 51.5)	12.4 (7.2 to 18.3)	−32.6 (−35.7 to −29.0)
Self-harm by other specified means	16	12	668.2 (577.4 to 883.1)	981.4 (771.2 to 1176.6)	46.9 (21.2 to 57.2)	19.1 (−2.0 to 27.6)	14.9 (12.7 to 19.5)	22.8 (17.7 to 27.6)	52.8 (27.9 to 62.4)	16.9 (−2.8 to 24.4)
Self-harm by firearm	14	13	827.2 (710.3 to 1071.4)	893.0 (700.7 to 1062.3)	7.9 (−8.3 to 16.3)	−16.9 (−29.1 to −10.7)	19.7 (16.3 to 24.6)	23.8 (18.5 to 27.9)	21.3 (3.2 to 29.5)	−13.2 (−26.1 to −7.3)
Pancreatic cancer	17	14	521.9 (511.0 to 532.4)	840.8 (816.1 to 865.1)	61.1 (55.8 to 67.0)	−2.5 (−5.7 to 1.1)	27.9 (27.4 to 28.5)	46.3 (44.9 to 47.6)	65.5 (60.3 to 71.3)	0.4 (−2.7 to 4.0)
Opioid use disorders	52	15	165.0 (149.9 to 202.4)	830.7 (393.3 to 924.6)	403.4 (145.1 to 482.9)	327.7 (109.9 to 395.1)	3.3 (3.0 to 4.1)	18.2 (8.5 to 20.3)	447.3 (165.2 to 533.9)	343.0 (112.5 to 413.2)
Chronic kidney disease due to diabetes mellitus	35	16	264.3 (240.1 to 287.6)	677.6 (620.8 to 727.7)	156.4 (146.1 to 166.5)	60.9 (54.9 to 66.7)	15.3 (13.6 to 17.0)	40.5 (36.3 to 45.0)	165.6 (153.6 to 177.2)	61.1 (55.3 to 66.6)
Hypertensive heart disease	26	17	387.3 (270.1 to 493.3)	669.9 (397.5 to 765.3)	73.0 (36.7 to 94.1)	9.3 (−15.5 to 23.5)	22.9 (15.7 to 29.2)	40.2 (25.5 to 47.8)	75.4 (51.6 to 88.9)	2.7 (−12.4 to 11.6)
Physical violence by firearm	15	18	797.4 (377.0 to 941.3)	659.8 (354.2 to 800.6)	−17.3 (−25.3 to 4.7)	−29.3 (−36.1 to −10.7)	14.4 (6.9 to 17.1)	12.4 (6.7 to 15.1)	−14.1 (−22.5 to 8.8)	−28.5 (−35.5 to −9.3)
Cirrhosis and other chronic liver diseases due to alcohol use	27	19	373.5 (351.2 to 396.9)	655.7 (606.6 to 703.5)	75.6 (67.6 to 83.3)	9.4 (4.7 to 13.9)	14.1 (13.3 to 14.9)	25.3 (23.5 to 27.0)	79.2 (71.7 to 86.5)	9.5 (5.3 to 13.7)
Other cardiovascular and circulatory diseases	18	20	518.2 (506.6 to 532.0)	636.9 (615.2 to 661.1)	22.9 (18.1 to 27.6)	−19.7 (−23.0 to −16.6)	30.7 (30.0 to 31.4)	42.2 (40.9 to 43.5)	37.6 (32.7 to 42.4)	−19.1 (−22.1 to −16.3)
Neonatal preterm birth complications	9	21	1036.0 (999.2 to 1074.2)	608.8 (572.3 to 638.0)	−41.2 (−44.9 to −37.9)	−39.9 (−43.6 to −36.5)	12.0 (11.5 to 12.4)	7.0 (6.6 to 7.4)	−41.3 (−44.9 to −37.9)	−39.9 (−43.7 to −36.5)
Endocrine, metabolic, blood, and immune disorders	37	22	248.1 (226.6 to 328.2)	587.7 (436.0 to 623.6)	136.8 (76.2 to 155.7)	60.3 (23.3 to 72.1)	8.5 (8.0 to 11.6)	26.2 (17.8 to 28.0)	208.7 (106.8 to 238.9)	89.1 (30.7 to 106.7)
Other neoplasms	24	23	411.2 (400.8 to 421.8)	570.6 (551.6 to 588.7)	38.8 (33.4 to 44.1)	−7.7 (−11.5 to −4.0)	16.6 (16.2 to 17.0)	28.6 (27.7 to 29.5)	72.6 (66.5 to 78.8)	5.3 (1.6 to 9.2)
Cirrhosis and other chronic liver diseases due to hepatitis C	30	24	304.9 (285.9 to 323.7)	513.5 (477.5 to 551.6)	68.4 (60.5 to 76.3)	7.0 (2.4 to 11.6)	10.8 (10.1 to 11.4)	18.9 (17.6 to 20.2)	75.4 (68.1 to 82.6)	8.4 (3.9 to 12.7)
Non-Hodgkin lymphoma	23	25	433.8 (371.6 to 448.0)	482.2 (459.8 to 564.9)	11.2 (4.7 to 37.3)	−30.0 (−34.3 to −11.7)	20.2 (17.9 to 20.8)	26.8 (25.7 to 29.6)	32.6 (26.5 to 50.3)	−19.5 (−23.3 to −7.3)

Abbreviations: UI, uncertainty interval; YLLs, years of life lost due to premature mortality.

Table 2. US Years Lived With Disability (YLDs) Rank, Rate, and Percentage Change for the 25 Leading Causes of Disability and Injury, 1990 and 2016

	YLD Rank		No. of YLDs, in Thousands (95% Uncertainty Interval)		% Change (95% Uncertainty Interval)	
Diseases and Injuries	1990	2016	1990	2016	YLDs	Age-Standardized YLD Rate
Low back pain	1	1	2461.1 (1732.3 to 3228.8)	3069.1 (2211.0 to 3989.6)	24.7 (10.9 to 39.6)	−12.4 (−22.3 to −1.9)
Major depressive disorder	2	2	1726.2 (1192.3 to 2330.5)	2193.0 (1507.6 to 2990.5)	27.0 (21.6 to 32.7)	0.1 (−4.1 to 3.7)
Diabetes mellitus	8	3	1040.2 (716.1 to 1450.3)	2142.9 (1496.1 to 2932.8)	106.0 (92.9 to 119.0)	29.6 (21.2 to 37.5)
Other musculoskeletal disorders	4	4	1573.5 (1076.0 to 2193.9)	2076.5 (1423.2 to 2843.1)	32.0 (22.5 to 42.0)	−2.3 (−8.7 to 4.3)
Migraine	3	5	1580.3 (1013.0 to 2205.4)	2010.1 (1296.4 to 2814.8)	27.2 (25.3 to 29.1)	−1.4 (−2.8 to −0.1)
Neck pain	6	6	1281.8 (880.4 to 1788.4)	1982.6 (1381.1 to 2704.7)	54.7 (39.2 to 73.3)	2.9 (−7.5 to 15.0)
Anxiety disorders	5	7	1341.7 (940.7 to 1813.1)	1755.0 (1229.6 to 2383.4)	30.8 (25.7 to 36.0)	0.6 (−3.2 to 4.5)
Opioid use disorders	7	8	1256.2 (892.8 to 1588.8)	1638.9 (1144.0 to 2097.4)	30.5 (23.3 to 37.4)	10.9 (4.2 to 17.2)
Age-related and other hearing loss	9	9	886.9 (599.0 to 1269.5)	1528.0 (1035.7 to 2157.1)	72.3 (67.3 to 78.3)	9.7 (6.6 to 13.4)
Falls	11	10	722.7 (487.8 to 1010.2)	1389.1 (960.0 to 1922.6)	92.2 (83.8 to 102.6)	22.2 (17.7 to 27.8)
Chronic obstructive pulmonary disease	12	11	674.4 (585.1 to 743.9)	1184.6 (1035.1 to 1307.3)	75.7 (66.8 to 84.9)	7.4 (2.0 to 12.9)
Osteoarthritis	14	12	573.5 (377.6 to 829.6)	1005.0 (659.5 to 1448.2)	75.2 (68.5 to 82.6)	7.9 (3.7 to 12.5)
Acne vulgaris	10	13	855.9 (573.6 to 1236.1)	992.6 (668.0 to 1441.4)	16.0 (14.3 to 17.8)	−1.5 (−3.0 to 0.2)
Dermatitis	13	14	659.8 (405.3 to 1032.9)	830.2 (515.1 to 1296.5)	25.8 (24.0 to 28.0)	1.2 (0.3 to 2.2)
Ischemic stroke	18	15	464.0 (321.3 to 607.7)	716.9 (500.9 to 912.7)	54.5 (41.5 to 64.9)	−3.5 (−11.2 to 2.9)
Schizophrenia	17	16	503.3 (365.0 to 627.7)	685.2 (506.9 to 847.9)	36.1 (32.7 to 39.5)	1.7 (−0.0 to 3.5)
Edentulism and severe tooth loss	19	17	458.0 (301.1 to 649.2)	662.2 (432.6 to 936.6)	44.6 (42.9 to 46.3)	−8.6 (−9.6 to −7.4)
Alcohol use disorders	15	18	558.2 (380.7 to 771.4)	633.9 (440.3 to 852.1)	13.6 (5.0 to 23.7)	−8.6 (−15.2 to −1.1)
Alzheimer disease and other dementias	23	19	360.8 (253.3 to 476.0)	597.6 (431.4 to 774.4)	65.6 (54.5 to 78.4)	−1.1 (−7.0 to 5.7)
Rheumatoid arthritis	25	20	350.2 (245.4 to 464.0)	592.4 (412.6 to 775.9)	69.1 (63.1 to 75.5)	10.1 (6.4 to 14.1)
Asthma	16	21	522.1 (342.5 to 744.9)	591.0 (393.8 to 832.4)	13.2 (7.5 to 19.2)	−12.4 (−17.1 to −7.6)
Other mental and substance use disorders	20	22	421.2 (288.4 to 601.1)	566.1 (390.0 to 804.1)	34.4 (32.9 to 35.9)	0.5 (−0.5 to 1.6)
Dysthymia	22	23	375.1 (254.3 to 548.2)	522.0 (352.7 to 753.3)	39.2 (31.7 to 47.2)	2.5 (−2.7 to 8.1)
Bipolar disorder	21	24	376.4 (236.7 to 545.7)	489.1 (309.6 to 703.2)	29.9 (27.9 to 32.1)	−0.4 (−1.9 to 1.0)
Psoriasis	24	25	351.9 (253.0 to 459.4)	488.2 (350.9 to 635.5)	38.7 (36.9 to 40.6)	1.9 (0.6 to 3.2)

falls had an increase of 22.2% in YLDs, opioid use disorders had an increase of 10.9% in YLDs, and asthma had a decline of 12.4% in age-standardized YLD rates. Other notable findings include an increase of 9.7% in age-standardized YLD rates of hearing loss due to aging and other causes.

US DALYs

Figure 1 shows the 25 leading causes of DALYs in 1990 and 2016 with their mean percentage change during the period. IHD and lung cancer were the leading causes of DALYs in both years, but the age-standardized rate declined between 1990 and 2016 by 49.7% for IHD and by 32.5% for lung cancer. The age-standardized DALY rate for chronic obstructive pulmonary disease (the third leading cause in 2016) increased by 5.5% between 1990 and 2016, and for diabetes (the fourth leading cause in 2016), it increased by 11%. Diabetes increased from the sixth leading cause in 1990 to the fourth in 2016, while low back pain declined from the third leading cause to the fifth. Three leading causes of DALYS had declines in age-standardized rates from 1990 to 2016: motor vehicle road injuries (by 35.0%), breast cancer (by 34.3%), and colorectal cancer (by 27.4%). Four leading causes of DALYS had increases in age-standardized rates from 1990 to 2016: opioid use disorders (by 47.9%), chronic kidney disease (by 44.3%), self-harm by other means (by 20.3%), and falls (by 19.0%).

US Risk Factor Estimates

Figure 2 shows the number of deaths and the percentage of DALYs from 17 risk factors in 2016. Diet, tobacco use, and high systolic blood pressure were the leading causes of deaths while tobacco use, high body mass index, and diet were the leading risk factors for DALYs. For example, dietary risks accounted for 529 299 deaths in 2016, with 83.9% of these deaths due to cardiovascular diseases, and the remainder due to a combination of neoplasms and diabetes, and to urogenital, blood, and endocrine diseases. Alcohol and drug use were the eighth leading cause of death and the fourth leading cause of DALYs. In 2016, each of the 6 following risks accounted for more than 5% of DALYs: tobacco consumption, high body mass index, diet, alcohol and drug use, high fasting plasma glucose levels, and high blood pressure.

Attribution of DALYs to Risk Factors

In 2016, 44.9% of total DALYs in the United States were attributable to risk factors. Behavioral risk factors accounted for the largest percentage of the attributable fraction of DALYs due to all causes (43.5%), followed by metabolic (22.7%), and environmental and occupational risks (3.7%) (Supplement 2).

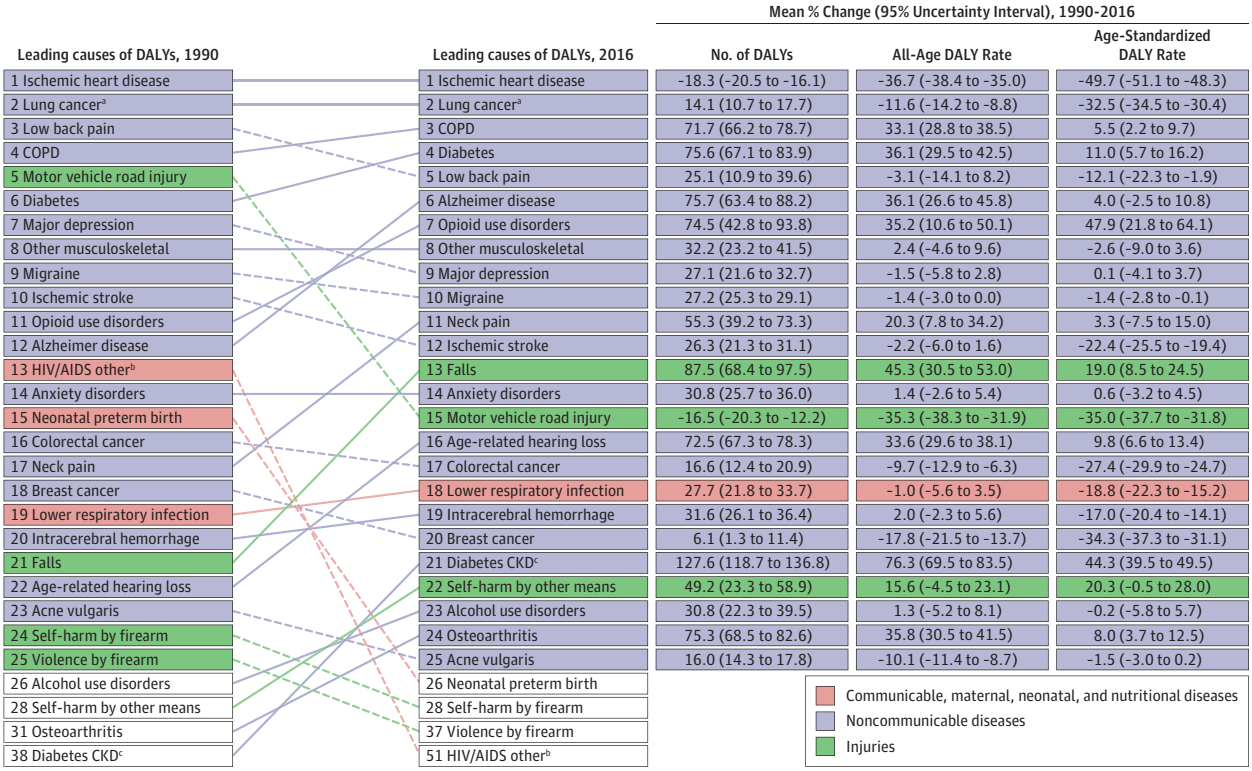
Individual State Data

GBD 2016 showed substantial variations in the burden of diseases, injuries, and risk factors by state. There was also a variation in trends by age, sex, and state (key findings and results of burden by state in Supplement 3).

Life Expectancy and HALE

Life expectancy and HALE at birth for both sexes combined for the United States, all 50 states, and for Washington, DC are shown in **Table 3**. Hawaii had the highest life expectancy at birth in 2016 (81.3 years [95% UI, 80.6 to 81.9]), while Mississippi had the lowest (74.7 years [95% UI, 73.5 to 76.1]; a 6.6-year difference). Other states with high life expectancy were California (80.9 years

Figure 1. Top 25 Causes of Disability-Adjusted Life-Years (DALYs) and % Change in Number of DALYs, All-Age DALYs, and Age-Standardized DALYs, 1990-2016



Dotted lines: a leading cause has decreased in rank between 1990 and 2016; solid lines, a cause has maintained or ascended to a higher ranking. Causes in white boxes were not among the top 25 in either 1990 or in 2016.
COPD, indicates chronic obstructive pulmonary disease.

[a] Includes tracheal, bronchus, and lung cancer.

[b] Indicates HIV/AIDS resulting in other diseases.

[c] Indicates chronic kidney disease (CKD) due to diabetes.

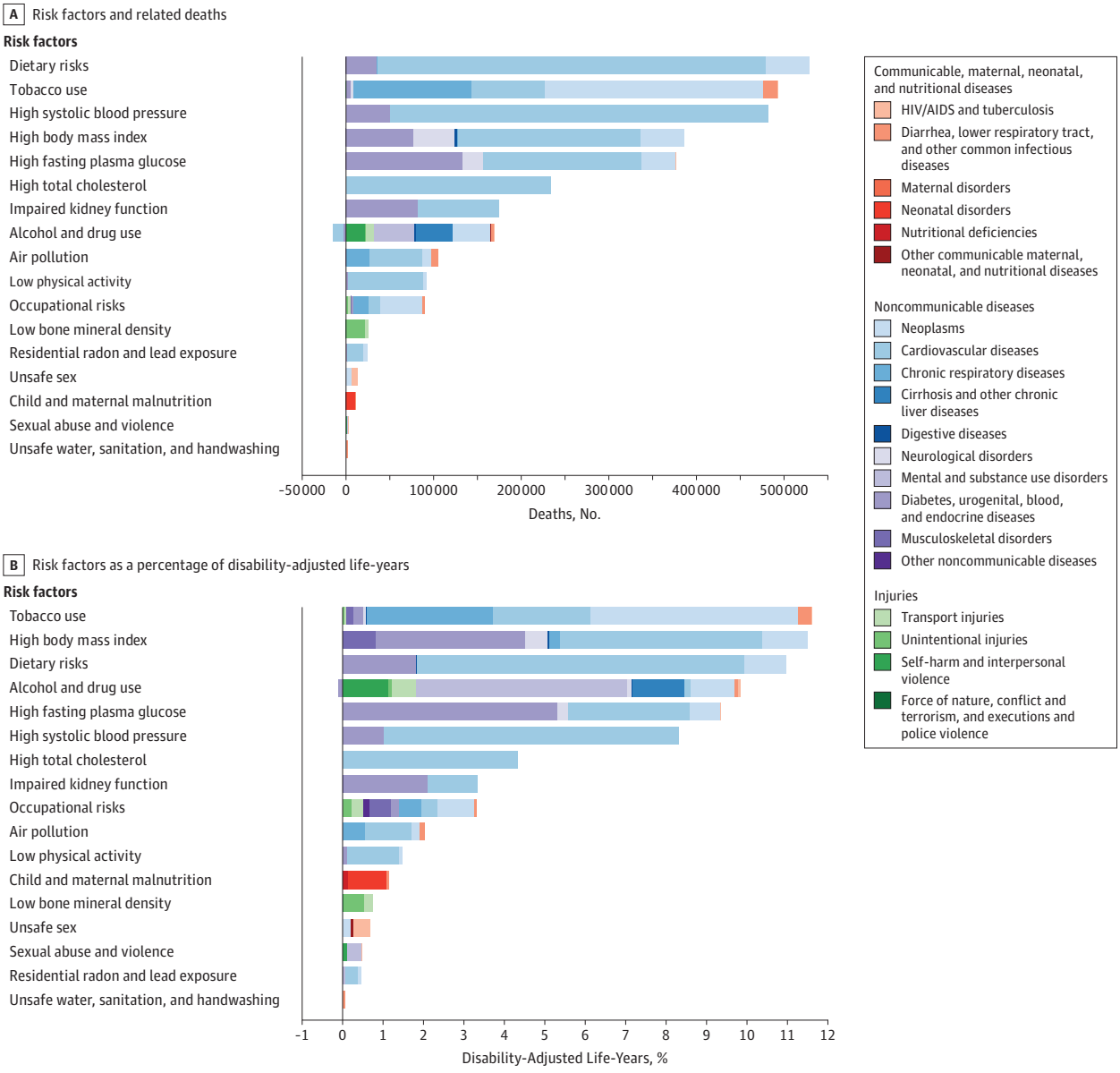
[95% UI, 79.9 to 81.9]), Connecticut (80.8 years [95% UI, 79.7 to 81.8]), Minnesota (80.8 years [95% UI, 80.0 to 81.6]), New York (80.5 years [95% UI, 79.4 to 81.6]), Massachusetts (80.4 years [95% UI, 79.6 to 81.1]), Colorado (80.2 years [95% UI, 79.4 to 80.9]), New Jersey (80.2 years [95% UI, 79.3 to 80.9]), and Washington (80.2 years [95% UI, 79.5 to 80.8]). Other states with low life expectancy were West Virigina (75.3 years [95% UI, 74.4 to 76.0]), Alabama (75.4 years [95% UI, 74.1 to 76.7]), Louisiana (75.6 years [95% UI, 74.9 to 76.4]), Oklahoma (75.7 years [95% UI, 75.0 to 76.4]), Arkansas (75.8 years [95% UI, 74.9 to 76.8]), and Kentucky (75.8 years [95% UI, 74.9 to 76.6]). In 2016, Minnesota had the highest HALE at birth with 70.3 years, while West Virginia had the lowest at 63.8 years, a 6.5-year difference. Only 2 states, Minnesota and Hawaii, had HALE values greater than 70.0 years at birth in 2016. In terms of life expectancy in 2016, only 9 states had life expectancy values greater than 80.0 years.

Male life expectancy and HALE at birth for the United States overall, for all states, and for Washington, DC, are shown in **Table 4**. Minnesota had the highest life expectancy in 2016 (78.7 years [95% UI, 77.5 to 79.8]) and HALE (69.1 years [95% UI, 66.3 to 71.9]), followed by California (life expectancy, 78.6 years [95% UI, 77.2 to 80.1]; and HALE, 68.6 years [95% UI, 65.5 to 71.6]). Mississippi had the lowest life expectancy for males in 2016 (71.8 years [95% UI, 70.1 to 73.8]) and ranked 49th for HALE (63.0 years [95% UI, 60.3 to 65.6]), while West Virginia had the lowest HALE (62.2 years [95% UI, 54.9 to 65.0]) and ranked 49th for life expectancy (72.7 years [95% UI, 71.5 to 73.9]).

Female life expectancy and HALE at birth for the United States, for all states, and for Washington, DC, are shown in **Table 5**. Hawaii had the highest life expectancy in 2016 (84.1 years [95% UI, 83.2 to 85.0) and HALE (71.9 years [95% UI, 68.3 to 75.1), followed by life expectancy for California (83.1 years [95% UI, 81.8 to 84.3]) and life expectancy for Connecticut (83.1 years [95% UI, 81.7 to 84.4]); the second highest HALE was for Minnesota (71.4 years [95% UI, 68.3 to 74.5]). Mississippi had the lowest life expectancy for females (77.7 years [95% UI, 76.1 to 79.6]), while West Virginia had the lowest HALE (65.5 years [95% UI, 61.9 to 68.5]).

Table 6 presents the age-standardized death rates, age-standardized YLL rates, and age-standardized YLD rates in 1990 and 2016 and their ranks by state. The 3 measurements varied widely between the states in 2016, ranging from 767.6 deaths per 100 000 in Mississippi to 465.8 deaths per 100 000 in Hawaii, from 17 775.9 YLLs per 100 000 in Mississippi to 9901.8 YLLs per 100 000 in Minnesota, and from 13 090.6 YLDs per 100 000 in West Virginia to 10 582.8 YLDs per 100 000 in Minnesota. A notable improvement was observed in Washington, DC (decreases from 1042.7 deaths per 100 000 to 603.3 deaths per 100 000, from 29 536.9 YLLS per 100 000 to 13 635.9 YLLs per 100 000, and from 12 230.8 YLDs per 100 000 to 11 421.1 YLDs per 100 000) and in California (decreases from 719.1 deaths per 100 000 to 491.7

Figure 2. Number of Deaths and Percentage of Disability-Adjusted Life-Years Related to the 17 Leading Risk Factors in the United States, 2016



Negative values (where bars extend left of zero) indicate a protective effect.

deaths per 100 000, from 15 903.4 YLLS per 100 000 to 9987.0 YLLs per 100 000, and from 11 170.5 YLDs per 100 000 to 10 990.4 YLDs per 100 000). Decreases in mortality and increases in morbidity were more apparent in Ohio (from 761.5 deaths per 100 000 to 644.1 deaths per 100 000, from 16 349.6 YLLs per 100 000 to 13 853.3 YLLs per 100 000, and from 12 009.0 YLDs per 100 000 to 12 334.7 YLDs per 100 000) and in Oklahoma (from 773.8 deaths per 100 000 to 725.3 deaths per 100 000, from 17 062.7 YLLs per 100 000 to 16 379.3 YLLs per 100 000, and from 12 036.5 YLDs per 100 000 to 12 549.7 YLDs per 100 000). The age-standardized death rates and age-standardized YLL rates declined in all states, but the level of decline for deaths ranged from 6.3% in Oklahoma to 42.1% in Washington, DC and the level of decline for YLLs ranged from 4.0% for Oklahoma to 53.8% for Washington, DC. Age-standardized YLD rates increased by 4.4% for West Virginia and declined by 6.6% for Washington, DC.

Decomposition of the Probability of Death by Age and State

The decomposition of change in the probability of death from birth to age 20 years, ages 20 to 55 years, and ages 55 to 90 years are shown in **Figure 3**, **Figure 4**, and **Figure 5**. For the United States and each state, these figures show the change in the probability of death from 1990 to 2016. In addition, these figures show changes in the probability of death over the interval due to changes in causes of death (GBD cause hierarchy level 2). The change in the probability of death from birth to age 20 years

Table 3. Life Expectancy and Healthy Life Expectancy at Birth for the United States, the 50 States, and Washington, DC, 1990 and 2016, Both Sexes

	Life Expectancy at Birth, y				Healthy Life Expectancy at Birth, y			
	1990		2016		1990		2016	
Location	Estimate (95% UI)	Rank	Estimate (95% UI)	Rank	Estimate (95% UI)	Rank	Estimate (95% UI)	Rank
United States	75.5 (75.4-75.5)		78.9 (78.7-79.0)		65.3 (62.4-67.9)		67.7 (64.5-70.5)	.
Alabama	73.7 (73.3-74.2)	47	75.4 (74.1-76.7)	49	63.7 (61.0-66.3)	48	64.6 (61.5-67.6)	48
Alaska	75.0 (74.5-75.6)	34	78.1 (76.9-79.4)	34	65.0 (62.2-67.6)	32	67.3 (64.2-70.3)	33
Arizona	76.1 (75.7-76.6)	22	79.5 (78.6-80.4)	19	65.4 (62.4-68.1)	24	67.7 (64.3-70.8)	27
Arkansas	74.4 (74.0-74.8)	43	75.8 (74.9-76.8)	45	64.5 (61.7-67.1)	40	65.5 (62.6-68.2)	44
California	75.9 (75.5-76.4)	24	80.9 (79.9-81.9)	2	66.1 (63.3-68.6)	19	69.9 (66.6-72.8)	3
Colorado	77.1 (76.7-77.4)	6	80.2 (79.4-80.9)	7	66.7 (63.7-69.4)	10	68.9 (65.7-71.8)	9
Connecticut	77.0 (76.6-77.5)	7	80.8 (79.7-81.8)	3	66.5 (63.5-69.1)	12	69.0 (65.7-72.1)	6
Delaware	74.9 (74.5-75.3)	37	78.6 (77.9-79.4)	28	64.6 (61.7-67.2)	39	67.2 (64.0-70.1)	34
Florida	76.0 (75.6-76.4)	23	79.6 (78.6-80.5)	15	65.5 (62.6-68.2)	23	67.9 (64.5-71.0)	24
Georgia	73.8 (73.2-74.3)	46	77.4 (76.2-78.4)	41	63.9 (61.0-66.5)	44	66.6 (63.5-69.5)	38
Hawaii	78.5 (78.2-78.9)	1	81.3 (80.6-81.9)	1	68.2 (65.2-70.9)	1	70.1 (66.9-73.0)	2
Idaho	77.0 (76.6-77.5)	8	79.1 (78.1-80.3)	23	66.4 (63.5-69.0)	13	67.9 (64.5-70.9)	25
Illinois	75.0 (74.7-75.4)	35	79.1 (78.4-79.8)	24	65.2 (62.5-67.7)	29	68.3 (65.3-71.0)	18
Indiana	75.4 (75.0-75.9)	27	77.2 (76.0-78.5)	42	65.1 (62.2-67.7)	31	66.0 (62.9-69.0)	42
Iowa	77.3 (77.0-77.7)	5	79.5 (78.6-80.4)	17	67.4 (64.6-70.0)	4	68.9 (65.8-71.7)	7
Kansas	76.8 (76.4-77.3)	12	78.5 (77.3-79.7)	30	66.8 (64.0-69.3)	8	67.8 (64.8-70.7)	26
Kentucky	74.4 (74.1-74.8)	42	75.8 (74.9-76.6)	46	63.7 (60.8-66.5)	45	64.3 (61.2-67.3)	50
Louisiana	73.3 (73.0-73.7)	49	75.6 (74.9-76.4)	48	63.2 (60.4-65.8)	50	65.0 (62.0-67.7)	46
Maine	76.3 (75.9-76.7)	21	79.0 (78.3-79.7)	26	66.1 (63.3-68.6)	17	68.0 (65.0-70.9)	21
Maryland	74.8 (74.5-75.2)	38	79.2 (78.5-79.9)	21	64.7 (61.8-67.3)	37	68.0 (65.0-70.7)	22
Massachusetts	76.7 (76.3-77.1)	15	80.4 (79.6-81.1)	6	66.1 (63.2-68.7)	18	68.9 (65.6-71.8)	8
Michigan	75.2 (74.9-75.5)	33	78.0 (77.3-78.7)	35	64.9 (62.0-67.6)	35	67.0 (63.9-69.9)	35
Minnesota	77.8 (77.4-78.2)	3	80.8 (80.0-81.6)	4	67.9 (65.1-70.4)	2	70.3 (67.4-73.0)	1
Mississippi	73.1 (72.6-73.6)	50	74.7 (73.5-76.1)	51	63.7 (61.0-66.1)	46	64.9 (62.1-67.6)	47
Missouri	75.3 (75.0-75.7)	31	77.4 (76.8-78.2)	40	65.3 (62.4-67.9)	28	66.5 (63.4-69.2)	39
Montana	76.4 (75.9-76.9)	18	78.9 (77.7-80.1)	27	66.0 (63.1-68.6)	20	67.7 (64.5-70.9)	28
Nebraska	76.9 (76.5-77.3)	10	79.4 (78.8-80.1)	20	67.0 (64.2-69.5)	7	68.8 (65.7-71.5)	12
Nevada	74.5 (74.1-74.9)	41	78.1 (77.3-79.0)	33	64.3 (61.5-67.0)	41	66.9 (63.9-69.7)	36
New Hampshire	76.7 (76.4-77.1)	14	79.9 (79.2-80.6)	11	66.2 (63.3-68.8)	16	68.5 (65.3-71.4)	14
New Jersey	75.4 (75.1-75.8)	26	80.2 (79.3-80.9)	8	65.3 (62.5-67.9)	25	68.8 (65.7-71.8)	10
New Mexico	75.9 (75.4-76.4)	25	77.8 (76.5-79.1)	38	65.3 (62.4-68.0)	26	66.3 (63.0-69.4)	40
New York	74.7 (74.3-75.2)	39	80.5 (79.4-81.6)	5	64.3 (61.4-67.0)	42	68.5 (65.1-71.6)	15
North Carolina	74.7 (74.4-75.1)	40	77.9 (77.3-78.6)	37	64.9 (62.2-67.5)	34	67.4 (64.4-70.1)	32
North Dakota	77.7 (77.2-78.2)	4	79.8 (78.8-80.9)	12	67.6 (64.7-70.1)	3	68.8 (65.6-71.8)	11
Ohio	75.4 (75.0-75.7)	29	77.5 (76.8-78.2)	39	64.9 (62.0-67.6)	33	66.1 (62.9-69.0)	41
Oklahoma	75.0 (74.6-75.4)	36	75.7 (75.0-76.4)	47	64.6 (61.7-67.2)	38	64.5 (61.4-67.4)	49
Oregon	76.3 (76.0-76.7)	19	79.5 (78.9-80.2)	16	66.0 (63.1-68.5)	21	68.4 (65.3-71.2)	17
Pennsylvania	75.4 (75.1-75.7)	28	78.5 (77.9-79.2)	29	64.7 (61.8-67.5)	36	66.8 (63.7-69.8)	37
Rhode Island	76.6 (76.1-77.0)	17	79.6 (78.6-80.7)	13	65.8 (62.8-68.5)	22	68.1 (64.9-71.2)	20
South Carolina	73.7 (73.2-74.2)	48	76.8 (75.7-77.8)	43	63.6 (60.8-66.2)	49	65.8 (62.7-68.7)	43
South Dakota	76.9 (76.4-77.4)	11	79.1 (78.1-80.1)	25	67.1 (64.3-69.6)	6	68.4 (65.3-71.4)	16
Tennessee	74.3 (74.0-74.7)	44	76.1 (75.5-77.0)	44	64.2 (61.4-66.8)	43	65.4 (62.3-68.2)	45
Texas	75.3 (74.8-75.6)	32	78.5 (77.8-79.3)	31	65.3 (62.5-67.9)	27	67.4 (64.4-70.3)	29
Utah	77.9 (77.6-78.3)	2	79.6 (79.0-80.3)	14	67.1 (64.1-69.8)	5	68.2 (64.9-71.1)	19
Vermont	76.6 (76.2-77.0)	16	79.9 (79.2-80.7)	10	66.4 (63.5-69.0)	14	69.0 (65.9-71.8)	5
Virginia	75.3 (75.0-75.7)	30	79.2 (78.5-79.9)	22	65.2 (62.3-67.8)	30	68.0 (64.9-70.8)	23
Washington	76.8 (76.5-77.2)	13	80.2 (79.5-80.8)	9	66.5 (63.7-69.1)	11	69.1 (65.8-71.9)	4
West Virginia	74.3 (74.0-74.7)	45	75.3 (74.4-76.0)	50	63.7 (60.7-66.3)	47	63.8 (60.7-66.7)	51
Wisconsin	76.9 (76.6-77.3)	9	79.5 (78.8-80.1)	18	66.8 (63.9-69.3)	9	68.6 (65.4-71.5)	13
Wyoming	76.3 (75.8-76.8)	20	78.4 (77.2-79.5)	32	66.2 (63.3-68.8)	15	67.4 (64.1-70.4)	30
Washington, DC	68.4 (67.7-69.0)	51	78.0 (76.8-79.1)	36	59.1 (56.6-61.6)	51	67.4 (64.4-70.3)	31

Abbreviation: UI, uncertainty interval.

Table 4. Life Expectancy and Healthy Life Expectancy at Birth for the United States, the 50 States, and Washington, DC, 1990 and 2016, Males

	Life Expectancy at Birth, y				Healthy Life Expectancy at Birth, y			
	1990		2016		1990		2016	
Location	Estimate (95% UI)	Rank	Estimate (95% UI)	Rank	Estimate (95% UI)	Rank	Estimate (95% UI)	Rank
United States	71.9 (71.8-72.1)		76.5 (76.2-76.7)		63.0 (60.5-65.2)		66.3 (63.5-68.8)	
Alabama	69.7 (69.1-70.4)	48	72.6 (70.8-74.5)	50	61.1 (58.6-63.4)	47	63.0 (60.1-65.8)	48
Alaska	71.9 (71.0-72.6)	31	75.9 (74.2-77.7)	33	63.0 (60.5-65.4)	29	66.2 (63.3-69.2)	31
Arizona	72.7 (72.1-73.3)	22	77.1 (75.9-78.5)	20	63.2 (60.6-65.6)	23	66.5 (63.3-69.4)	27
Arkansas	70.6 (70.0-71.2)	44	73.3 (71.8-74.7)	45	62.0 (59.6-64.4)	41	64.1 (61.5-66.5)	44
California	72.6 (72.0-73.3)	23	78.6 (77.2-80.1)	2	63.9 (61.5-66.2)	20	68.6 (65.5-71.6)	2
Colorado	73.9 (73.4-74.5)	6	78.1 (77.0-79.1)	6	64.9 (62.3-67.3)	6	67.8 (65.0-70.6)	5
Connecticut	73.7 (73.1-74.4)	10	78.4 (76.7-79.9)	4	64.3 (61.6-66.7)	15	67.6 (64.5-70.6)	8
Delaware	71.7 (71.2-72.3)	33	76.2 (75.1-77.4)	29	62.6 (60.0-64.9)	35	65.8 (63.0-68.5)	33
Florida	72.3 (71.7-72.9)	25	77.0 (75.4-78.3)	22	63.1 (60.6-65.4)	27	66.4 (63.3-69.4)	28
Georgia	69.8 (69.1-70.6)	46	74.8 (73.1-76.3)	42	61.3 (58.8-63.5)	45	65.2 (62.2-68.0)	38
Hawaii	75.5 (75.0-76.1)	1	78.4 (77.4-79.4)	3	66.3 (63.7-68.8)	1	68.3 (65.3-71.1)	3
Idaho	74.0 (73.3-74.7)	5	77.2 (75.6-78.8)	16	64.6 (62.0-67.0)	12	67.0 (63.8-70.0)	19
Illinois	71.5 (70.9-72.0)	37	76.6 (75.5-77.6)	26	62.9 (60.5-65.1)	31	66.8 (64.0-69.4)	21
Indiana	72.0 (71.4-72.7)	28	74.8 (73.1-76.5)	41	63.1 (60.5-65.4)	28	64.8 (61.9-67.7)	40
Iowa	73.9 (73.3-74.5)	7	77.2 (76.0-78.3)	18	65.2 (62.7-67.5)	5	67.6 (64.8-70.1)	7
Kansas	73.5 (72.8-74.1)	13	76.1 (74.4-77.9)	31	64.7 (62.3-67.0)	10	66.5 (63.7-69.5)	25
Kentucky	70.8 (70.3-71.3)	42	73.2 (72.0-74.4)	47	61.3 (58.7-63.8)	44	62.8 (59.9-65.7)	50
Louisiana	69.4 (68.8-69.9)	49	72.9 (71.8-74.0)	48	60.7 (58.3-62.9)	50	63.3 (60.7-65.8)	47
Maine	72.9 (72.4-73.5)	21	76.5 (75.4-77.5)	27	64.0 (61.5-66.2)	18	66.5 (63.6-69.2)	26
Maryland	71.4 (70.9-71.9)	38	76.8 (75.8-77.7)	24	62.5 (60.1-64.8)	37	66.7 (63.9-69.4)	22
Massachusetts	73.3 (72.8-73.9)	14	77.9 (76.9-78.9)	9	63.9 (61.3-66.4)	21	67.4 (64.3-70.1)	12
Michigan	71.9 (71.4-72.3)	32	75.6 (74.7-76.7)	35	62.8 (60.3-65.1)	34	65.6 (62.8-68.4)	35
Minnesota	74.5 (74.0-75.0)	3	78.7 (77.5-79.8)	1	65.9 (63.5-68.1)	2	69.1 (66.3-71.9)	1
Mississippi	69.0 (68.3-69.7)	50	71.8 (70.1-73.8)	51	60.9 (58.6-63.0)	49	63.0 (60.3-65.6)	49
Missouri	71.6 (71.0-72.1)	34	74.9 (73.9-76.0)	40	62.9 (60.5-65.2)	30	65.1 (62.2-67.6)	39
Montana	73.2 (72.5-73.9)	18	76.8 (75.1-78.7)	23	64.1 (61.4-66.3)	17	66.6 (63.5-69.8)	23
Nebraska	73.6 (73.0-74.0)	12	77.2 (76.1-78.1)	19	64.9 (62.5-67.1)	7	67.5 (64.7-70.1)	9
Nevada	71.3 (70.7-71.8)	39	75.9 (74.6-77.1)	34	62.3 (59.8-64.6)	40	65.7 (62.9-68.4)	34
New Hampshire	73.6 (73.1-74.2)	11	77.7 (76.6-78.7)	11	64.2 (61.6-66.5)	16	67.1 (64.3-70.0)	17
New Jersey	72.2 (71.6-72.7)	26	77.8 (76.6-78.9)	10	63.1 (60.6-65.5)	24	67.3 (64.4-70.1)	14
New Mexico	72.4 (71.7-73.2)	24	75.0 (73.1-76.9)	39	63.1 (60.5-65.6)	25	64.7 (61.4-67.9)	42
New York	70.9 (70.3-71.6)	40	78.1 (76.6-79.7)	5	61.7 (59.2-64.1)	42	67.1 (64.0-70.1)	16
North Carolina	70.9 (70.4-71.3)	41	75.4 (74.4-76.4)	36	62.3 (59.9-64.5)	39	65.9 (63.2-68.4)	32
North Dakota	74.2 (73.6-74.9)	4	77.2 (75.7-78.8)	15	65.5 (63.0-67.7)	3	67.4 (64.3-70.2)	13
Ohio	72.1 (71.7-72.6)	27	75.1 (74.0-76.1)	37	62.9 (60.3-65.1)	33	64.8 (61.8-67.4)	41
Oklahoma	71.6 (71.0-72.1)	35	73.2 (72.1-74.4)	46	62.5 (60.0-64.8)	36	63.3 (60.5-66.1)	46
Oregon	73.2 (72.7-73.7)	17	77.4 (76.5-78.3)	13	63.9 (61.3-66.3)	19	67.2 (64.5-69.9)	15
Pennsylvania	71.9 (71.5-72.4)	29	76.0 (74.9-77.0)	32	62.4 (59.8-64.8)	38	65.3 (62.5-68.0)	37
Rhode Island	73.1 (72.5-73.7)	19	77.2 (75.6-78.5)	17	63.6 (61.0-66.1)	22	66.6 (63.7-69.6)	24
South Carolina	69.8 (69.0-70.5)	47	74.2 (72.5-75.9)	43	61.0 (58.5-63.3)	48	64.3 (61.4-67.1)	43
South Dakota	73.2 (72.6-73.9)	16	76.7 (75.1-78.2)	25	64.7 (62.2-66.9)	11	67.1 (64.2-70.0)	18
Tennessee	70.5 (70.0-71.0)	45	73.5 (72.5-74.7)	44	61.6 (59.2-63.9)	43	63.8 (61.0-66.4)	45
Texas	71.5 (70.9-72.1)	36	76.1 (75.0-77.3)	30	62.9 (60.4-65.2)	32	66.2 (63.5-69.0)	30
Utah	75.0 (74.5-75.5)	2	77.9 (77.0-78.9)	8	65.5 (62.8-67.9)	4	67.5 (64.7-70.2)	10
Vermont	73.3 (72.8-73.9)	15	77.6 (76.4-78.7)	12	64.3 (61.7-66.6)	14	67.7 (64.8-70.3)	6
Virginia	71.9 (71.4-72.3)	30	77.0 (76.0-78.0)	21	63.1 (60.6-65.4)	26	67.0 (64.1-69.6)	20
Washington	73.8 (73.3-74.3)	8	78.1 (77.1-79.1)	7	64.8 (62.2-67.0)	9	68.1 (65.1-70.8)	4
West Virginia	70.6 (70.1-71.2)	43	72.7 (71.5-73.9)	49	61.2 (58.6-63.5)	46	62.2 (59.4-65.0)	51
Wisconsin	73.7 (73.2-74.2)	9	77.3 (76.2-78.2)	14	64.8 (62.3-67.1)	8	67.4 (64.6-70.1)	11
Wyoming	73.1 (72.4-73.8)	20	76.2 (74.6-77.8)	28	64.3 (61.7-66.7)	13	66.3 (63.2-69.3)	29
Washington, DC	62.3 (61.4-63.3)	51	75.1 (73.4-76.6)	38	54.5 (52.3-56.8)	51	65.5 (62.6-68.3)	36

Abbreviation: UI, uncertainty interval.

Table 5. Life Expectancy and Healthy Life Expectancy at Birth for the United States, the 50 States, and Washington, DC, 1990 and 2016, Females

	Life Expectancy at Birth, y				Healthy Life Expectancy at Birth, y			
	1990		2016		1990		2016	
Location	Estimate (95% UI)	Rank	Estimate (95% UI)	Rank	Estimate (95% UI)	Rank	Estimate (95% UI)	Rank
United States	78.9 (78.8-79.0)		81.2 (81.0-81.5)		67.5 (64.2-70.4)		69.0 (65.5-72.1)	
Alabama	77.6 (77.0-78.3)	46	78.1 (76.6-79.7)	49	66.2 (63.0-69.1)	46	66.2 (62.7-69.5)	48
Alaska	78.8 (78.0-79.5)	28	80.5 (78.9-82.1)	36	67.4 (64.2-70.3)	31	68.5 (65.0-72.0)	34
Arizona	79.7 (79.1-80.2)	19	81.9 (80.7-83.0)	15	67.7 (64.4-70.8)	25	69.0 (65.2-72.4)	25
Arkansas	78.2 (77.6-78.8)	39	78.5 (77.3-79.8)	45	66.9 (63.8-69.8)	34	67.1 (63.7-70.1)	44
California	79.2 (78.7-79.9)	25	83.1 (81.8-84.3)	2	68.2 (65.1-71.0)	16	71.1 (67.7-74.3)	3
Colorado	80.0 (79.6-80.5)	11	82.3 (81.3-83.2)	9	68.4 (65.1-71.5)	11	69.9 (66.3-73.1)	13
Connecticut	80.1 (79.5-80.7)	9	83.1 (81.7-84.4)	3	68.4 (65.2-71.4)	12	70.4 (66.8-73.9)	4
Delaware	77.9 (77.4-78.4)	45	81.0 (80.0-82.0)	29	66.4 (63.2-69.4)	44	68.6 (65.2-71.8)	32
Florida	79.7 (79.2-80.2)	16	82.2 (81.1-83.4)	11	67.9 (64.6-71.0)	22	69.4 (65.8-72.8)	21
Georgia	77.5 (77.0-78.2)	47	79.8 (78.2-81.4)	41	66.4 (63.2-69.3)	45	68.1 (64.6-71.1)	38
Hawaii	81.8 (81.3-82.3)	1	84.1 (83.2-85.0)	1	70.2 (67.0-73.1)	1	71.9 (68.3-75.1)	1
Idaho	80.1 (79.5-80.8)	8	81.2 (79.8-82.5)	26	68.3 (65.0-71.4)	13	68.8 (65.2-72.1)	28
Illinois	78.4 (77.9-78.8)	36	81.5 (80.6-82.3)	23	67.5 (64.4-70.3)	28	69.8 (66.4-72.7)	16
Indiana	78.7 (78.1-79.3)	31	79.6 (77.9-81.2)	42	67.1 (63.8-70.1)	33	67.3 (63.8-70.6)	43
Iowa	80.6 (80.1-81.1)	6	81.9 (80.8-83.1)	16	69.5 (66.4-72.4)	5	70.2 (66.8-73.4)	9
Kansas	80.1 (79.5-80.7)	10	80.9 (79.1-82.5)	30	68.7 (65.5-71.7)	7	69.1 (65.7-72.4)	23
Kentucky	78.1 (77.5-78.6)	42	78.4 (77.3-79.4)	46	66.1 (62.8-69.1)	48	65.8 (62.2-69.1)	50
Louisiana	77.1 (76.6-77.7)	50	78.4 (77.4-79.4)	47	65.7 (62.5-68.7)	50	66.6 (63.2-69.8)	47
Maine	79.5 (79.0-80.0)	22	81.5 (80.7-82.4)	22	68.2 (65.0-71.1)	15	69.6 (66.3-72.8)	18
Maryland	78.1 (77.7-78.6)	40	81.5 (80.6-82.4)	21	66.7 (63.5-69.7)	38	69.2 (65.8-72.4)	22
Massachusetts	79.8 (79.2-80.2)	15	82.7 (81.8-83.6)	6	68.0 (64.8-71.0)	19	70.3 (66.8-73.7)	5
Michigan	78.4 (78.0-78.8)	34	80.4 (79.5-81.4)	37	66.9 (63.7-69.9)	35	68.3 (64.9-71.4)	36
Minnesota	80.9 (80.5-81.4)	3	82.9 (81.8-83.9)	4	69.9 (66.7-72.7)	2	71.4 (68.3-74.5)	2
Mississippi	77.2 (76.6-77.8)	49	77.7 (76.1-79.6)	51	66.5 (63.5-69.2)	43	66.8 (63.7-70.0)	46
Missouri	79.0 (78.5-79.4)	27	79.9 (79.0-80.9)	39	67.6 (64.3-70.5)	26	67.9 (64.5-71.0)	40
Montana	79.7 (79.0-80.4)	17	81.0 (79.4-82.5)	28	68.1 (64.9-71.1)	18	68.8 (65.2-72.2)	29
Nebraska	80.1 (79.6-80.6)	7	81.8 (80.8-82.7)	17	68.9 (65.7-71.9)	6	70.0 (66.6-73.0)	11
Nevada	78.0 (77.5-78.5)	43	80.5 (79.5-81.7)	35	66.6 (63.4-69.5)	41	68.2 (64.7-71.4)	37
New Hampshire	79.7 (79.2-80.1)	18	82.2 (81.4-83.0)	13	68.0 (64.7-71.0)	20	69.9 (66.2-73.0)	12
New Jersey	78.5 (78.1-79.0)	32	82.4 (81.4-83.5)	8	67.4 (64.2-70.3)	30	70.2 (66.8-73.4)	8
New Mexico	79.4 (78.8-80.1)	24	80.7 (79.1-82.3)	33	67.5 (64.2-70.6)	27	67.9 (64.3-71.5)	39
New York	78.3 (77.8-78.9)	38	82.7 (81.3-84.2)	5	66.7 (63.4-69.8)	39	69.7 (65.9-73.3)	17
North Carolina	78.4 (77.9-78.8)	35	80.3 (79.5-81.2)	38	67.4 (64.3-70.2)	29	68.8 (65.4-71.7)	30
North Dakota	81.3 (80.6-82.0)	2	82.6 (81.3-83.9)	7	69.8 (66.5-72.7)	3	70.3 (66.8-73.6)	7
Ohio	78.5 (78.1-79.0)	33	79.9 (79.1-80.7)	40	66.8 (63.5-69.9)	37	67.4 (63.9-70.5)	41
Oklahoma	78.3 (77.9-78.8)	37	78.2 (77.2-79.2)	48	66.6 (63.3-69.6)	42	65.8 (62.3-68.9)	49
Oregon	79.4 (79.0-79.9)	23	81.7 (80.8-82.6)	19	68.0 (64.8-70.9)	21	69.5 (66.3-72.7)	19
Pennsylvania	78.7 (78.3-79.1)	29	81.0 (80.2-81.9)	27	66.9 (63.6-70.0)	36	68.4 (64.7-71.6)	35
Rhode Island	79.7 (79.1-80.3)	20	82.0 (80.7-83.3)	14	67.8 (64.5-70.8)	23	69.4 (65.8-72.9)	20
South Carolina	77.4 (76.8-78.1)	48	79.4 (77.7-80.9)	43	66.1 (62.9-69.0)	49	67.3 (63.7-70.6)	42
South Dakota	80.6 (80.0-81.2)	5	81.6 (80.1-83.0)	20	69.5 (66.3-72.4)	4	69.8 (66.4-72.9)	15
Tennessee	78.1 (77.7-78.6)	41	78.8 (77.8-79.7)	44	66.7 (63.6-69.6)	40	66.9 (63.4-70.1)	45
Texas	79.0 (78.5-79.5)	26	80.9 (79.8-81.9)	31	67.7 (64.5-70.6)	24	68.6 (65.2-71.9)	31
Utah	80.8 (80.3-81.2)	4	81.4 (80.5-82.2)	24	68.7 (65.2-71.7)	8	68.8 (65.3-71.9)	27
Vermont	79.8 (79.3-80.3)	13	82.2 (81.2-83.3)	12	68.5 (65.3-71.4)	10	70.3 (67.0-73.5)	6
Virginia	78.7 (78.3-79.2)	30	81.3 (80.4-82.1)	25	67.2 (64.0-70.2)	32	69.0 (65.6-72.1)	26
Washington	79.8 (79.3-80.3)	14	82.3 (81.4-83.2)	10	68.3 (65.0-71.2)	14	70.0 (66.4-73.2)	10
West Virginia	78.0 (77.5-78.5)	44	77.9 (76.8-78.9)	50	66.2 (62.9-69.1)	47	65.5 (61.9-68.5)	51
Wisconsin	80.0 (79.6-80.5)	12	81.7 (80.8-82.5)	18	68.6 (65.4-71.5)	9	69.8 (66.3-73.0)	14
Wyoming	79.6 (78.8-80.3)	21	80.8 (79.3-82.2)	32	68.1 (64.8-71.1)	17	68.6 (64.9-71.9)	33
Washington, DC	74.6 (73.9-75.3)	51	80.6 (79.1-82.2)	34	63.9 (60.8-66.6)	51	69.1 (65.6-72.6)	24

Abbreviation: UI, uncertainty interval.

Table 6. Age-Standardized Rates of Death, Years of Life Lost Due to Premature Mortality, and Years Lived With Disability for the United States, the 50 States, and Washington, DC, 1990 and 2016, Both Sexes

Location	Age-Standardized Death Rate, per 100 000				Age-Standardized YLL Rate, per 100 000				Age-Standardized YLD Rate, per 100 000			
	1990		2016		1990		2016		1990		2016	
	Rate (95% UI)	Rank	Rate (95% UI)	Rank	Rate (95% UI)	Rank	Rate (95% UI)	Rank	Rate (95% UI)	Rank	Rate (95% UI)	Rank
United States	745.2 (740.6-749.8)	.	578.0 (569.4-587.1)	.	16518.3 (16410.3-16632.5)	.	12257.5 (12064.5-12452.7)	.	11644.0 (8660.1-14968.1)	.	11717.7 (8722.3-15059.7)	.
Alabama	831.2 (801.9-859.4)	4	740.2 (667.5-814.5)	3	19179.9 (18491.3-19854.3)	5	16851.4 (15167.5-18586.1)	3	11872.3 (8849.6-15244.8)	16	12249.1 (9128.0-15787.0)	8
Alaska	765.6 (733.3-800.7)	18	610.4 (546.2-670.7)	17	17185.2 (16461.9-17961.4)	15	13423.3 (11995.1-14839.1)	16	11485.6 (8574.6-14812.8)	32	11517.8 (8588.1-14848.5)	34
Arizona	699.6 (677.5-721.6)	35	539.3 (498.6-578.3)	38	15824.6 (15295.3-16355.4)	30	12045.3 (11125.9-12950.2)	27	12141.3 (9021.7-15624.5)	7	12242.7 (9112.6-15658.6)	9
Arkansas	793.7 (767.6-819.7)	11	715.0 (666.7-764.7)	7	18151.2 (17514.5-18753.1)	7	16176.9 (15046.7-17365.8)	6	11579.7 (8597.3-14853.0)	31	11604.9 (8631.1-14935.5)	33
California	719.1 (695.1-743.6)	28	491.7 (449.2-535.5)	50	15903.4 (15362.8-16464.3)	29	9987.0 (9086.3-10951.1)	50	11170.5 (8300.7-14399.6)	44	10990.4 (8165.7-14201.2)	50
Colorado	668.6 (649.8-688.9)	44	522.5 (491.0-558.1)	45	14269.3 (13850.9-14687.4)	44	10784.5 (10092.7-11584.1)	43	11414.0 (8489.6-14634.4)	37	11495.6 (8556.5-14785.0)	35
Connecticut	669.4 (643.8-694.7)	43	496.8 (453.0-546.0)	49	14278.3 (13719.3-14847.9)	43	10012.6 (9102.2-11076.8)	49	11755.8 (8752.1-15078.4)	23	11932.9 (8912.6-15337.7)	18
Delaware	787.6 (763.2-813.1)	12	582.9 (548.1-618.9)	24	17107.6 (16506.0-17672.8)	17	12784.4 (11939.3-13641.8)	21	12024.5 (8970.5-15407.6)	10	12107.8 (9049.6-15513.8)	13
Florida	694.0 (672.2-715.4)	38	532.0 (493.7-575.3)	41	16414.6 (15886.2-16933.9)	23	11938.8 (11058.2-12931.1)	29	11738.6 (8743.2-15142.9)	24	11960.0 (8900.6-15356.1)	17
Georgia	829.2 (800.1-861.5)	6	652.7 (599.1-716.1)	11	19071.7 (18365.8-19834.3)	6	13945.8 (12753.4-15335.5)	13	11633.4 (8641.4-15016.0)	27	11652.8 (8615.6-15024.4)	29
Hawaii	586.4 (569.0-604.2)	51	465.8 (439.8-494.3)	51	12739.0 (12341.0-13138.5)	51	10138.2 (9539.8-10788.7)	48	11028.6 (8216.7-14261.0)	46	11090.8 (8247.3-14405.3)	47
Idaho	668.4 (641.3-695.0)	45	571.4 (518.1-627.1)	27	14390.0 (13800.5-14964.0)	42	11790.6 (10629.1-13011.9)	32	11706.8 (8715.5-15042.3)	25	11730.9 (8810.9-15210.6)	24
Illinois	769.4 (748.5-788.9)	17	570.6 (536.2-606.5)	29	17158.5 (16639.6-17617.6)	16	11878.6 (11150.8-12673.1)	31	11242.3 (8353.2-14443.3)	43	11300.2 (8468.7-14597.7)	41
Indiana	758.6 (730.9-785.9)	23	657.7 (593.8-726.2)	10	16326.9 (15731.9-16937.8)	27	14178.4 (12757.5-15633.6)	11	11765.1 (8748.8-15122.3)	22	12154.0 (9119.3-15588.8)	10
Iowa	658.8 (638.2-678.7)	47	556.0 (516.4-602.1)	32	13773.1 (13328.3-14213.5)	47	11215.6 (10359.6-12152.9)	38	10821.7 (8041.9-13938.7)	49	11030.0 (8153.3-14290.3)	48
Kansas	678.8 (652.5-703.3)	41	595.5 (537.8-658.1)	21	14549.4 (13970.3-15080.3)	37	12617.6 (11369.9-13923.7)	23	11104.9 (8213.9-14332.6)	45	11307.9 (8434.3-14600.1)	40
Kentucky	813.3 (791.9-834.9)	8	729.2 (685.1-777.9)	4	17708.0 (17218.0-18214.3)	12	16047.8 (15015.7-17183.9)	7	12574.6 (9336.3-16086.0)	1	13044.8 (9781.6-16602.9)	2
Louisiana	853.4 (830.0-875.8)	3	725.3 (685.7-766.9)	5	19771.1 (19200.4-20286.8)	3	16507.6 (15563.3-17560.9)	4	12087.5 (9002.4-15500.5)	8	12110.8 (9002.1-15512.5)	12
Maine	723.1 (701.8-745.1)	27	580.7 (544.3-615.7)	25	14859.5 (14414.0-15343.1)	34	11980.0 (11222.5-12759.6)	28	11452.5 (8488.9-14749.9)	35	11468.3 (8522.1-14767.5)	36
Maryland	782.6 (762.4-803.3)	13	558.7 (527.5-592.7)	30	17261.4 (16806.5-17738.0)	14	11931.6 (11229.7-12656.4)	30	11796.6 (8772.9-15103.9)	20	11647.5 (8744.3-14997.4)	30
Massachusetts	694.9 (671.5-716.4)	36	519.3 (488.6-556.2)	46	14452.9 (13954.6-14918.3)	40	10316.3 (9656.5-11084.0)	46	11931.8 (8856.0-15314.6)	14	11720.4 (8695.3-15038.6)	25
Michigan	761.8 (741.1-782.4)	19	618.1 (585.5-653.0)	15	16839.8 (16397.3-17288.8)	19	13266.5 (12512.2-14079.3)	18	11797.4 (8764.9-15181.8)	19	11800.8 (8835.1-15095.9)	21
Minnesota	639.5 (619.0-660.4)	48	499.8 (462.1-536.5)	48	13167.4 (12740.2-13617.9)	49	9901.8 (9123.5-10668.9)	51	10673.6 (7960.0-13781.8)	51	10582.8 (7913.2-13640.3)	51
Mississippi	856.9 (824.8-889.2)	2	767.6 (695.6-838.6)	1	20205.8 (19404.7-20980.9)	2	17775.9 (16050.5-19528.3)	1	11265.3 (8375.4-14543.6)	41	11256.4 (8444.8-14484.4)	44
Missouri	751.4 (730.8-772.6)	26	642.6 (607.0-678.5)	13	16661.6 (16188.1-17157.3)	21	13999.7 (13137.5-14844.2)	12	11481.7 (8537.8-14818.3)	33	11814.3 (8845.4-15220.7)	20
Montana	694.6 (666.4-723.2)	37	571.2 (517.1-628.5)	28	15261.1 (14641.3-15880.3)	32	12459.2 (11201.1-13768.0)	25	11583.2 (8605.5-14871.2)	30	11655.8 (8712.6-15,013.9)	28
Nebraska	676.7 (656.8-696.9)	42	557.9 (525.3-591.3)	31	14395.9 (13955.8-14862.7)	41	11400.3 (10658.6-12162.4)	35	10949.8 (8148.9-14134.8)	47	11020.1 (8203.9-14215.1)	49
Nevada	805.1 (779.9-828.2)	10	613.3 (569.4-656.4)	16	17800.5 (17209.9-18358.0)	10	12987.3 (12032.1-14000.4)	20	11867.6 (8782.2-15278.2)	17	12002.9 (8964.4-15428.2)	16
New Hampshire	701.7 (682.8-723.6)	33	537.8 (505.5-573.0)	40	14213.2 (13787.3-14665.6)	46	10855.5 (10153.7-11633.6)	41	11828.2 (8772.6-15147.7)	18	11742.3 (8757.5-15080.9)	22
New Jersey	755.3 (733.9-777.1)	24	526.2 (490.1-567.1)	43	16378.6 (15892.3-16881.0)	24	10605.8 (9827.2-11477.5)	44	11588.6 (8620.3-14901.0)	29	11643.9 (8632.8-15014.8)	31
New Mexico	705.1 (679.7-734.8)	31	608.6 (549.2-669.2)	18	16272.8 (15656.8-16969.0)	28	14218.1 (12673.0-15771.9)	10	11991.5 (8920.9-15393.5)	13	12401.4 (9248.5-15907.9)	5
New York	773.1 (746.1-798.3)	16	508.6 (460.7-560.7)	47	17741.2 (17099.9-18338.1)	11	10279.3 (9279.1-11391.9)	47	12168.8 (9045.5-15688.9)	5	12254.7 (9140.2-15737.0)	7
North Carolina	779.3 (757.7-800.6)	14	622.3 (589.0-656.4)	14	17628.0 (17133.8-18121.3)	13	13366.2 (12620.1-14165.6)	17	11288.9 (8412.6-14533.0)	39	11291.6 (8386.1-14568.8)	42
North Dakota	637.8 (610.9-663.9)	49	525.7 (479.7-574.5)	44	13499.1 (12938.2-14041.9)	48	11431.6 (10389.4-12556.6)	34	10892.6 (8091.1-14040.5)	48	11259.2 (8397.6-14507.9)	43
Ohio	761.5 (741.1-781.2)	20	644.1 (608.7-679.9)	12	16349.6 (15901.8-16791.9)	26	13853.3 (13037.7-14672.9)	14	12009.0 (8918.1-15405.3)	12	12334.7 (9199.1-15781.7)	6
Oklahoma	773.8 (752.5-796.3)	15	725.3 (686.4-763.9)	6	17062.7 (16588.6-17559.8)	18	16379.3 (15465.9-17299.4)	5	12036.5 (8968.0-15488.4)	9	12549.7 (9358.4-16161.4)	3
Oregon	708.9 (690.0-728.4)	29	552.8 (521.3-582.0)	36	15122.3 (14700.2-15557.7)	33	11300.8 (10622.8-11944.4)	37	11692.7 (8722.5-14973.6)	26	11658.6 (8651.6-14977.2)	26

(continued)

Table 6. Age-Standardized Rates of Death, Years of Life Lost Due to Premature Mortality, and Years Lived With Disability for the United States, the 50 States, and Washington, DC, 1990 and 2016, Both Sexes (continued)

Location	Age-Standardized Death Rate, per 100 000				Age-Standardized YLL Rate, per 100 000				Age-Standardized YLD Rate, per 100 000			
	1990		2016		1990		2016		1990		2016	
	Rate (95% UI)	Rank	Rate (95% UI)	Rank	Rate (95% UI)	Rank	Rate (95% UI)	Rank	Rate (95% UI)	Rank	Rate (95% UI)	Rank
Pennsylvania	759.6 (739.9-778.7)	22	595.0 (561.4-629.4)	22	16362.8 (15935.8-16789.4)	25	12619.1 (11865.9-13353.8)	22	12223.0 (9041.3-15663.8)	4	12506.3 (9378.1-15987.4)	4
Rhode Island	701.0 (675.1-726.0)	34	548.5 (500.7-599.7)	37	14693.5 (14149.1-15250.3)	36	11190.8 (10192.0-12316.8)	40	12149.4 (9011.5-15564.7)	6	12124.4 (9041.0-15542.0)	11
South Carolina	829.3 (799.1-859.1)	5	676.8 (625.2-739.6)	9	19287.6 (18569.1-19989.3)	4	14846.7 (13657.2-16269.8)	9	12015.1 (8923.8-15458.1)	11	12085.6 (8987.0-15537.6)	14
South Dakota	663.2 (636.1-688.3)	46	555.4 (510.0-602.8)	33	14810.8 (14208.8-15379.4)	35	12432.3 (11400.0-13504.4)	26	10786.0 (8017.8-13942.2)	50	11106.1 (8289.1-14338.2)	46
Tennessee	807.9 (786.6-830.0)	9	709.4 (664.7-747.1)	8	18039.2 (17547.9-18571.2)	8	15635.0 (14629.3-16495.9)	8	11910.5 (8870.9-15244.3)	15	12064.7 (8975.6-15485.5)	15
Texas	753.5 (731.3-776.8)	25	599.4 (561.2-636.5)	20	16787.5 (16275.7-17342.1)	20	12553.3 (11703.8-13359.6)	24	11385.2 (8500.3-14726.4)	38	11633.7 (8651.6-14997.3)	32
Utah	632.8 (615.4-651.7)	50	555.1 (522.5-587.6)	34	13049.2 (12687.1-13428.1)	50	11192.0 (10486.6-11895.6)	39	11783.7 (8714.8-15138.7)	21	11909.5 (8867.1-15402.9)	19
Vermont	704.8 (681.6-728.4)	32	539.2 (505.5-577.5)	39	14457.3 (13950.0-14957.0)	39	10792.5 (10081.0-11626.2)	42	11434.0 (8503.5-14647.4)	36	11243.2 (8344.9-14511.7)	45
Virginia	760.3 (742.2-780.7)	21	572.2 (538.3-607.0)	26	16447.1 (16036.4-16911.3)	22	11683.4 (10979.4-12452.2)	33	11589.6 (8609.5-15005.0)	28	11657.5 (8674.1-15008.9)	27
Washington	682.2 (662.4-701.4)	40	526.9 (497.1-559.2)	42	14501.1 (14069.7-14941.5)	38	10587.5 (9942.6-11291.9)	45	11462.0 (8489.0-14725.3)	34	11419.6 (8471.8-14708.2)	38
West Virginia	820.0 (795.0-844.6)	7	750.6 (709.1-799.7)	2	17864.8 (17332.1-18407.8)	9	16969.9 (15984.9-18167.6)	2	12540.2 (9410.5-16014.5)	2	13090.6 (9811.0-16681.6)	1
Wisconsin	683.0 (663.3-703.5)	39	554.9 (525.5-589.2)	35	14247.1 (13842.1-14679.3)	45	11389.6 (10740.6-12101.9)	36	11261.6 (8330.2-14507.8)	42	11310.4 (8374.7-14641.7)	39
Wyoming	705.4 (677.6-734.0)	30	589.8 (539.2-645.8)	23	15278.9 (14678.7-15918.9)	31	13089.2 (11911.6-14432.0)	19	11268.4 (8384.2-14546.1)	40	11740.6 (8719.4-15186.0)	23
Washington, DC	1042.7 (1002.8-1079.5)	1	603.3 (549.7-660.8)	19	29536.9 (28403.2-30634.1)	1	13635.9 (12369.7-15028.7)	15	12230.8 (9121.5-15731.3)	3	11421.1 (8482.8-14709.1)	37

Abbreviations: UI, uncertainty interval; YLD, years lived with disability; YLL, years of life lost due to premature mortality.

declined in all states. The most pronounced declines were in South Carolina, Georgia, Alaska, and New York, at higher than a 0.85-point decrease in probability (Figure 3). In contrast, Maine had the lowest decline of 0.32-point probability. In the United States as a whole, there was a decline of 0.70, which was associated with neonatal disorders, other noncommunicable diseases (including congenital), and a large contribution from injuries, with slight increases from mental and substance use disorders (Figure 3).

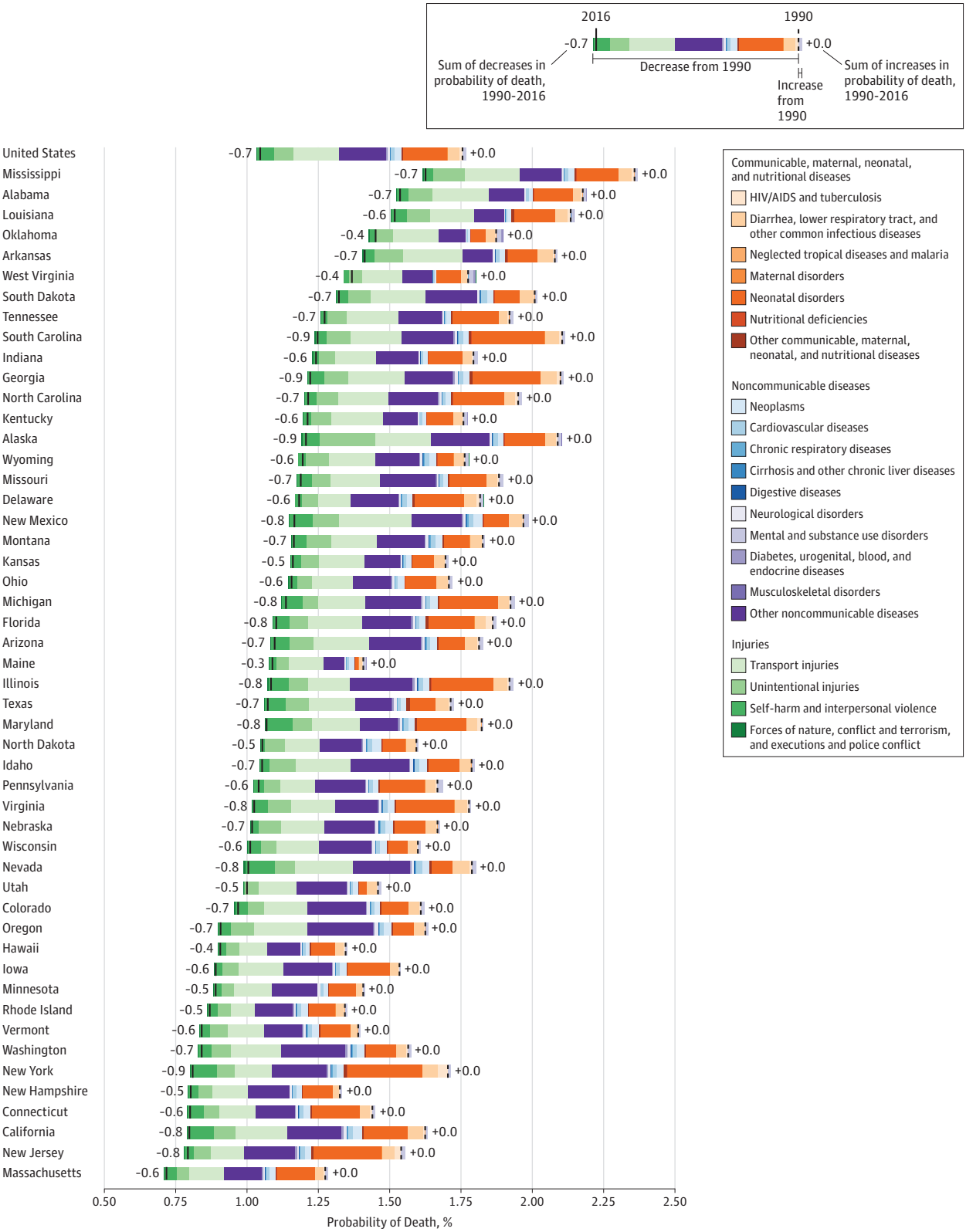
The largest reductions in probability of death for ages 20 to 55 years were observed in New York (3.5) and California (2.5) and the highest increases were observed in West Virginia (2.6) and Oklahoma (2.0) (Figure 4). In 21 states, the probability of death has actually increased from 1990 to 2016, but of these, only 5 showed an increase of greater than 10% (Kentucky, Oklahoma, New Mexico, West Virginia, and Wyoming). Conversely, 31 states and Washington, DC have seen decreases in the probability of death among adults aged 20 to 55 years over the same period, but only in 15 states was that decrease more than 10% (New York, California, Illinois, New Jersey, Georgia, Maryland, Florida, Nevada, Texas, Virginia, Connecticut, North Carolina, Massachusetts, Washington, South Carolina). Decreases in the probability of death in US states were influenced by declines in HIV/AIDS across all state groups, as well as declines in road injuries and neoplasms, while increases in probability of death were influenced by increased burden of drug use disorders, alcohol use disorders, and chronic kidney disease, among others (Figure 4).

All states experienced a considerable reduction in probabilities of death for ages 55 to 90 years, largely associated with reductions in the probability of dying from cardiovascular diseases (Figure 5). The highest point decline was observed in California at 12.6 points, compared with lowest decline of 3.5 points for Mississippi. These declines were somewhat offset by increases in the death rates associated with cirrhosis and other liver disease, neurological disorders, and mental and substance use disorders in all states. Hawaii was the only state in which the probability of death was less than 65% for ages 55 to 90 years. Other notable findings include the variation in the decline of probability of death between the 3 age groups, with an 8.3-point decline of probability of death for ages 55 to 90 years, a 1-point decline for ages 20 to 55 years, and a 0.7-point decline for ages 0 to 20 years.

YLLs Overall, by State, and for Washington, DC

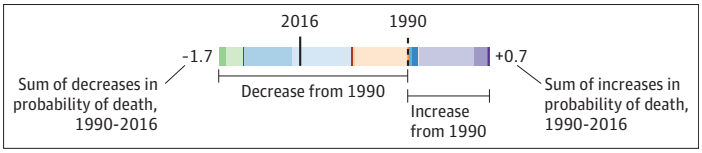
The age-standardized YLL rates for the United States, all states, and Washington, DC in 2016 for the top 20 causes are grouped by 3 levels of significance (**Figure 6**) indicating significantly below the mean, indistinguishable from the mean, and significantly higher than the mean. The heat map shows a clear pattern of performance ranging from Minnesota to Mississippi, with some clear patterns of exception for some causes. For example, Colorado had a YLL rate from self-harm significantly above the mean (760), while Washington, DC had a YLL rate significantly lower than the mean (306). Mississippi, West Virginia, Alabama, Oklahoma, Kentucky, Tennessee, South Carolina, Indiana, Missouri, and Ohio had YLLs significantly higher than the mean with a few exceptions of causes that were indistinguishable from the mean. Other notable findings were that Louisiana had YLLs for

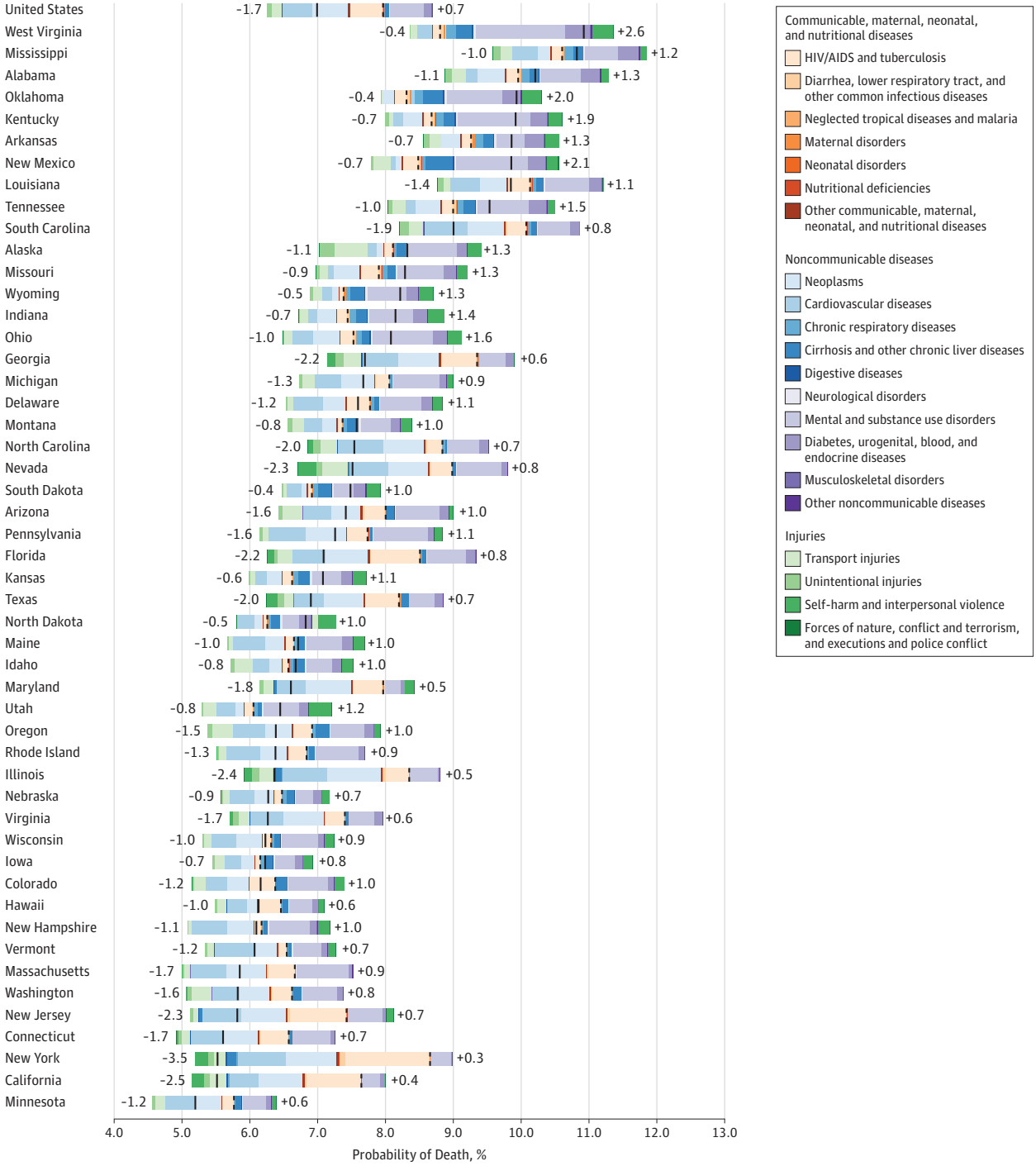
Figure 3. Change in the Probability of Death Between Birth and Age 20 Years, 1990-2016, Reported Showing Major Causes of Death for the United States Overall and the 50 States



States are listed in descending order according to probability of death in 2016. Data for Washington, DC, were not included in this analysis.

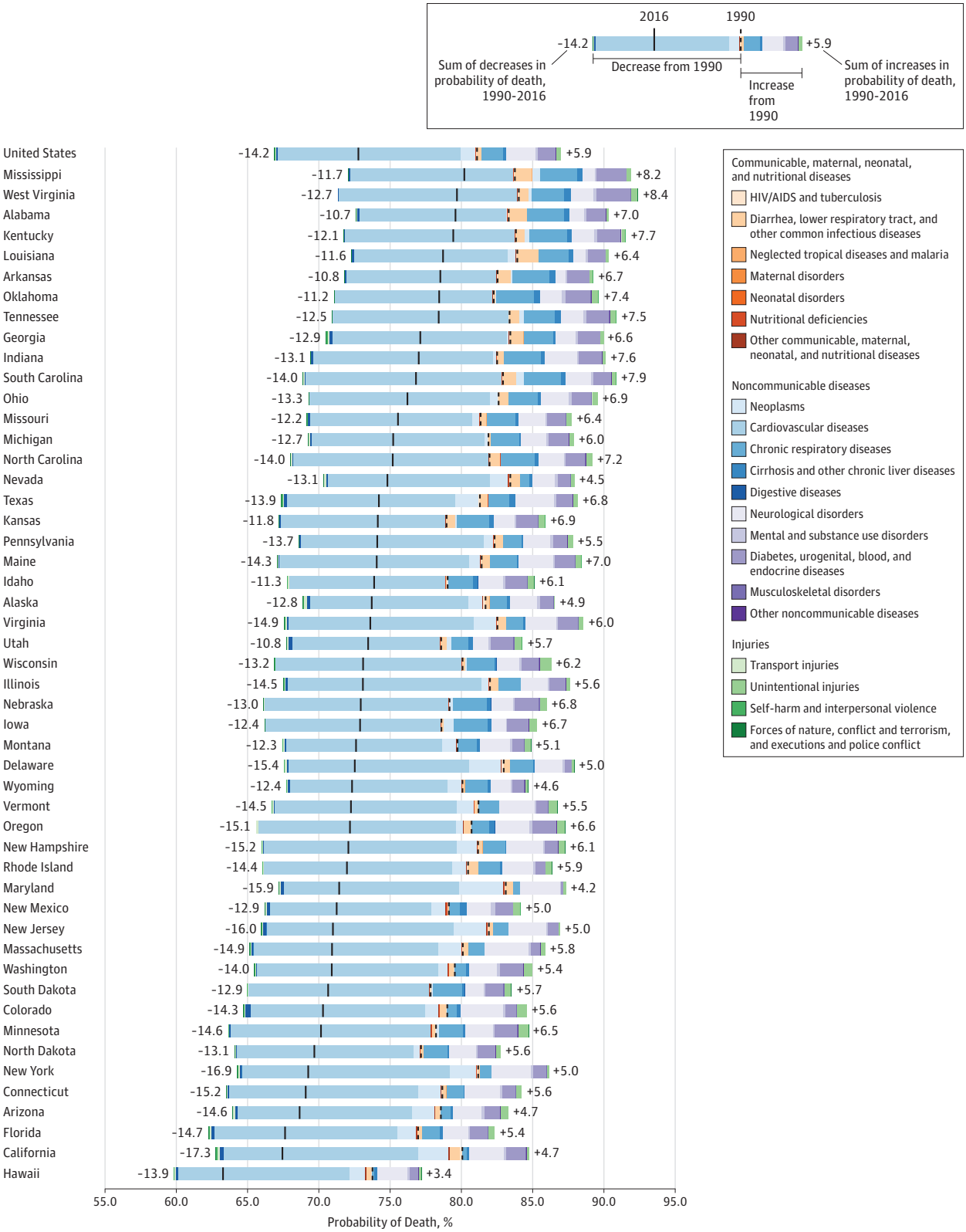
Figure 4. Change in the Probability of Death Between Ages 20 and 55 Years, 1990-2016, Reported Showing Major Causes of Death for the United States Overall and the 50 States





States are listed in descending order according to probability of death in 2016. Data for Washington, DC, were not included in this analysis.

Figure 5. Change in the Probability of Death Between Ages 55 and 90 Years, 1990-2016, Reported Showing Major Causes of Death for the United States Overall and the 50 States



States are listed in descending order according to probability of death in 2016. Data for Washington, DC, were not included in this analysis.

Figure 6. Age-Standardized Rates of Years of Life Lost per 100 000 Persons for the 20 Leading Causes of Years of Life Lost in 2016 for the United States Overall, the 50 States, and the District of Columbia

Age-standardized rates of years of life lost per 100 000
Significantly lower than mean US rate | Indistinguishable from mean US rate | Significantly higher than mean US rate

	Ischemic heart disease	Tracheal, bronchus, and lung cancer	Road injuries	Self-harm	Cerebrovascular disease	Chronic obstructive pulmonary disease	Drug use disorders	Alzheimer disease and other dementias	Colon and rectum cancer	Interpersonal violence	Lower respiratory tract infections	Diabetes	Congenital birth defects	Neonatal preterm birth complications	Chronic kidney disease	Breast cancer	Pancreatic cancer	Cardiomyopathy and myocarditis	Cirrhosis and other chronic liver diseases due to alcohol use	Endocrine, metabolic, blood, and immune disorders
United States	1651	784	602	564	503	501	451	360	323	318	308	299	293	280	270	247	184	161	156	155
Minnesota	1044	693	452	534	421	413	258	349	279	163	203	234	302	198	203	214	177	113	112	124
California	1346	519	450	429	421	365	353	324	281	288	219	238	249	179	213	223	164	155	159	121
Connecticut	1231	671	422	405	349	351	480	365	284	217	255	214	202	267	196	222	183	133	131	124
Hawaii	1261	607	437	563	461	229	289	385	315	180	275	225	249	272	216	194	188	199	134	130
New York	1699	667	344	374	303	331	374	349	287	267	287	241	227	229	201	234	180	123	111	125
Massachusetts	1280	747	344	466	375	374	555	380	294	197	301	201	202	241	222	227	185	116	137	121
Washington	1196	671	422	598	448	467	455	335	274	200	202	271	250	183	172	216	175	134	152	158
New Jersey	1479	659	395	383	388	334	481	326	322	289	288	294	209	235	239	255	185	156	126	157
Colorado	1135	511	529	760	405	532	439	389	264	224	245	204	271	225	180	206	162	113	174	123
Vermont	1251	784	520	647	360	510	346	364	288	185	190	254	291	208	168	215	168	117	118	192
New Hampshire	1273	789	501	644	361	491	473	395	290	169	212	256	247	224	186	230	182	117	126	172
Rhode Island	1555	846	352	491	361	433	514	396	311	187	238	249	254	335	196	238	178	110	148	158
Utah	1225	398	489	979	512	437	410	323	267	176	312	347	316	194	221	236	166	118	138	170
Iowa	1632	825	585	564	468	555	223	306	333	155	269	246	290	170	189	227	179	127	110	152
Oregon	1159	748	485	694	504	534	413	355	294	184	199	307	252	203	187	235	186	115	179	183
Wisconsin	1443	755	561	607	465	447	410	351	292	217	251	239	297	298	236	228	191	145	121	145
Nebraska	1321	794	683	569	494	596	199	344	334	212	257	274	311	216	237	241	184	122	130	162
North Dakota	1473	731	770	638	444	422	161	368	311	183	231	303	369	271	211	216	177	118	128	186
Virginia	1516	788	504	533	521	446	330	395	314	277	352	274	278	331	294	262	185	160	134	140
Idaho	1362	666	696	772	493	591	313	366	280	183	244	291	312	219	205	231	183	117	155	166
Illinois	1654	799	462	457	487	444	437	355	327	377	333	267	285	350	288	250	188	173	124	135
Maryland	1564	718	490	694	461	358	216	398	311	429	337	289	258	403	259	262	188	192	118	167
Florida	1472	775	679	566	434	459	503	332	305	347	221	299	289	282	235	236	174	181	167	161
Maine	1346	922	614	631	436	575	420	375	312	169	258	289	349	301	229	217	185	139	138	197
Arizona	1408	626	677	735	411	502	564	324	283	345	264	307	331	236	224	233	169	138	204	159
South Dakota	1574	765	871	681	463	507	198	309	326	213	280	297	447	274	205	250	178	121	159	166
Montana	1373	717	918	879	444	604	328	382	300	222	257	269	335	229	182	227	175	118	172	159
Texas	1741	713	721	529	574	505	346	413	321	321	349	318	313	253	323	239	173	156	191	149
Kansas	1628	860	719	652	532	635	328	351	326	261	322	296	348	291	272	237	189	142	148	162
Pennsylvania	1678	827	558	587	477	460	708	358	348	330	323	304	276	333	276	257	199	172	140	157
Delaware	1604	880	644	581	453	471	529	328	311	309	288	315	301	468	278	265	185	184	148	214
Nevada	1721	802	561	759	526	632	599	323	354	332	398	213	275	211	292	250	190	186	212	118
Wyoming	1514	670	923	876	436	702	431	336	310	239	314	269	325	275	192	221	176	127	188	172
Michigan	1995	863	553	634	489	530	589	377	334	394	285	320	302	370	276	259	197	196	140	167
North Carolina	1695	898	732	552	647	583	447	359	323	347	388	328	321	398	334	263	187	178	156	187
Alaska	1443	875	514	915	530	530	519	361	341	328	251	301	266	197	201	224	189	158	183	156
District of Columbia	1984	797	253	306	440	283	428	399	380	625	318	332	204	432	353	320	232	214	190	186
Ohio	1882	966	566	602	561	631	691	393	359	337	338	383	348	364	305	272	199	181	159	204
Georgia	1861	884	726	541	717	614	389	360	373	392	448	335	297	372	377	271	187	191	150	161
Missouri	2072	1003	748	673	560	645	542	389	348	395	346	310	322	314	304	268	195	173	143	158
Indiana	1965	1010	684	697	581	703	502	425	357	353	340	360	355	347	335	265	197	193	151	187
New Mexico	1466	563	950	903	425	543	852	388	334	404	303	417	329	258	273	240	176	136	328	212
South Carolina	1970	965	936	593	760	640	452	380	363	407	397	361	297	340	367	289	201	205	192	187
Tennessee	2324	1124	864	685	723	710	603	385	382	412	435	372	367	319	325	277	204	182	186	203
Kentucky	2323	1306	875	710	624	851	812	438	428	313	435	383	365	254	367	269	203	163	182	187
Arkansas	2492	1169	1020	788	768	758	330	338	401	434	473	408	373	268	373	285	205	173	177	172
Oklahoma	2569	1059	973	791	677	829	611	416	385	395	401	403	418	333	325	279	196	170	219	179
Louisiana	2327	1055	931	636	758	592	551	312	438	624	488	446	375	429	489	302	230	220	185	204
Alabama	2326	1143	1086	690	847	774	495	366	418	515	498	438	393	438	414	305	221	192	201	180
West Virginia	2317	1215	946	781	636	860	982	393	434	341	415	526	391	301	397	293	193	178	212	273
Mississippi	2664	1210	1242	638	850	736	399	344	452	552	517	452	416	418	499	323	232	216	190	181

Boxes are colored green if significantly less than the US rate ($P<.05$), red if significantly more than the US rate ($P<.05$), and yellow if not significantly different from the US rate ($P\geq.05$).

all causes higher than that of the US mean except for Alzheimer diseases and other dementias that was significantly higher than the mean, while Georgia had only 1 cause, drug use disorder, that was significantly above the US mean.

Observed to Expected YLLs Overall, by State, and for Washington, DC

The ratio of observed YLLs to those expected, based on the SDI for the 10 leading causes of YLLs for the United States overall, each individual state, and Washington, DC, are shown in **Figure 7**. For example, in Alabama, the stroke ratio was 1, indicating that the observed rates are similar to what would be expected given the state's SDI, whereas diabetes is observed at 2.85 times more than expected. For the United States overall, the ratios of observed YLLs to those expected were 0.75 for IHD, 1.11 for lung cancer, 1.88 for chronic obstructive pulmonary disease, 0.61 for stroke, 1.19 for road injury, 1.14 for Alzheimer disease, 0.98 for self-harm, 7.17 for drugs, 0.84 for colorectal cancer, and 1.27 for lower respiratory tract infection. In general, most states performed better on IHD and stroke but worse for chronic obstructive pulmonary disease and drug use disorders. Colorado had the best performance for IHD at 0.58 observed to expected ratio of YLLs, while West Virginia had the highest observed to expected ratio of YLLs for drug use disorders at 14.38. Other notable findings are the high rates in Washington, DC for interpersonal violence, drug use disorders, HIV/AIDS, and chronic kidney disease, given that Washington, DC has the highest SDI rank in the US and in the world.

Observed to Expected YLDs by State and for Washington, DC

The ratio of the YLDs observed to those expected based on SDI for the 10 leading causes of YLDs for the United States overall, the 50 states, and Washington, DC are shown in **Figure 8**. Minnesota had lower observed YLDs from low back and neck pain (0.63 ratio) and migraine (0.79), but higher YLD rates from drug use disorders (2.32). Most states had lower than expected YLDs from low back and neck pain but higher rates for drug use disorders and other musculoskeletal disorders. A notable finding is the higher than expected rates of YLDs from depression, anxiety, and skin diseases, and lower than expected rates from falls in most states.

Observed to Expected DALYs by State and for Washington, DC

The ratio of DALYs observed to those expected based on SDI for the 10 leading causes of DALYs for the United States overall, the 50 states, and Washington, DC are shown in **Figure 9**. All states had lower than expected DALY rates from IHD except for Alabama, Arkansas, Kentucky, Mississippi, Oklahoma, Tennessee, and West Virginia. The highest observed rates of DALYs compared with expected from drug use disorders were in West Virginia at 7.77 and in Kentucky at 7.31. Other notable findings are the higher than expected rates of drug use disorders and interpersonal violence in Washington, DC, and the lower than expected rates of lung cancer in California.

Changes in Age-Standardized Summary Exposure Values

The age-standardized percent changes in summary exposure values for the top 10 risk factors from 1990 to 2016 are shown in **Table 7**. High fasting plasma glucose increased by 76% and high body mass index by 53.2%, while smoking declined by 42.8% during the same time period in the United States. There were clear patterns in these variations by states. High fasting plasma glucose increased in all states; the increase ranged from 127.2% in Mississippi to 1.7% in Pennsylvania. Drug use increased in all states except Arkansas, Maryland, and Oregon. Other notable findings include reductions in high systolic blood pressure, high total cholesterol levels, and diet low in whole grains in all states.

Leading Risk Factors for DALYs by State and for Washington, DC

The rank of risk factors by state in 2016 are shown in **Figure 10**. Tobacco smoking was the leading risk factor for DALYs in the United States and in 33 states. Alcohol and drug use were the leading risk for DALYs for 7 states and Washington, DC, while high BMI was the leading cause for 10 states (California, Connecticut, Hawaii, Illinois, Maryland, New Jersey, New York, North Dakota, Texas, and Virginia). Another notable finding is that diet was the third leading cause of DALYs in the United States overall but the second in 20 states.

Discussion

This study provides a comprehensive report on the burden of disease and its patterns in the United States and the individual 50 states from 1990 to 2016 and reveals wide disparities in burden of disease at the state level. Moreover, these findings show distinct trends in different age bands and demonstrate that improvement in some health outcomes, such as IHD, lung cancer, and neonatal preterm complications, are balanced by rising death rates from drug use disorders, chronic obstructive pulmonary disease, self-harm, chronic kidney disease, cirrhosis, and hypertensive heart disease. Summary measures, such as life expectancy, that do not differentiate the trends in different age groups mask the heterogeneous directions for US health status by age and state. Above and beyond the drivers of divergent trends, the study reveals that there has been far greater progress in reducing the burden of some major causes of YLLs, such as IHD and lung cancer, but no progress in addressing some of the leading causes of YLDs such as mental health disorders and musculoskeletal disorders. These findings should be used to examine the causes of health variations and to plan, develop, and implement programs and policies to improve health overall and eliminate disparities in the United States.

Mortality reversals in 21 states for adults ages 20 to 55 years are strongly linked to the burden of substance use disorders, cirrhosis, and self-harm, and this study shows that the trends for some of these conditions differ considerably across different states. Case and Deaton have called some of these conditions "deaths of despair" and argued that they are linked to the social and economic status of white US adults.[3] A wide range of interventions have been proposed to address substance abuse, cirrhosis, and self-harm. For substance abuse,[28,47-49] prevention programs should account for the root causes of substance use, the socioeconomic factors involved, and causes of relapses during treatment.[50,51] Physicians have a major role to play in addiction control by counseling their patients who are on medication for pain control.[52-54] For cirrhosis, intervention strategies to treat hepatitis C and decrease excessive alcohol consumption are important. For self-harm, the most promising approaches relate to decreasing the case-fatality rate from suicide attempts by restricting access to lethal means; in the United States, a large share of suicides are due to firearms.[55,56] While multiple strategies are available for dealing with these problems, they have not until very

Figure 7. Ratio of Observed Years of Life Lost (YLLs) to Expected YLLs Based on the Sodiodemographic Index (SDI) for the United States Overall, the 50 States, and the District of Columbia in 2016 for the 10 Leading Causes in Each Jurisdiction

	1	2	3	4	5	6	7	8	9	10
United States	IHD (0.75)	Lung cancer (1.11)	COPD (1.88)	Stroke (0.61)	Road injury (1.19)	Alzheimer (1.14)	Self-harm (0.98)	Drugs (7.17)	Colorectal cancer (0.84)	LRI (1.27)
Alabama	IHD (1.05)	Lung cancer (1.71)	Stroke (1.0)	COPD (2.9)	Road injury (1.97)	Self-harm (1.16)	LRI (1.98)	Diabetes (2.85)	Colorectal cancer (1.13)	Alzheimer (1.17)
Alaska	IHD (0.5)	Lung cancer (1.0)	Self-harm (1.62)	COPD (1.42)	Drugs (8.52)	Stroke (0.46)	Road injury (1.04)	Alcohol (6.59)	Colorectal cancer (0.72)	Diabetes (1.64)
Arizona	IHD (0.65)	Lung cancer (0.95)	COPD (1.9)	Self-harm (1.21)	Road injury (1.22)	Stroke (0.5)	Drugs (8.59)	Alzheimer (1.12)	Diabetes (1.95)	Colorectal cancer (0.77)
Arkansas	IHD (1.12)	Lung cancer (1.76)	COPD (2.76)	Stroke (0.9)	Road injury (1.77)	Self-harm (1.28)	LRI (1.82)	Alzheimer (1.15)	Colorectal cancer (1.1)	Diabetes (2.54)
California	IHD (0.58)	Lung cancer (0.67)	Stroke (0.47)	COPD (1.26)	Alzheimer (0.93)	Self-harm (0.77)	Road injury (0.92)	Drugs (6.0)	Colorectal cancer (0.68)	Diabetes (1.52)
Colorado	IHD (0.58)	Self-harm (1.39)	COPD (1.9)	Lung cancer (0.66)	Road injury (1.14)	Alzheimer (1.01)	Stroke (0.46)	Drugs (7.28)	Colorectal cancer (0.63)	LRI (0.97)
Connecticut	IHD (0.64)	Lung cancer (0.98)	Alzheimer (1.3)	COPD (1.57)	Stroke (0.5)	Drugs (7.54)	Colorectal cancer (0.76)	Self-harm (0.76)	Road injury (0.94)	LRI (1.29)
Delaware	IHD (0.8)	Lung cancer (1.34)	COPD (2.02)	Stroke (0.61)	Road injury (1.29)	Self-harm (1.01)	Alzheimer (1.1)	Drugs (8.12)	Colorectal cancer (0.86)	Diabetes (2.36)
District of Columbia	IHD (0.75)	Lung cancer (0.84)	Violence (8.19)	Alzheimer (0.93)	Stroke (0.46)	Drugs (7.72)	HTN HD (3.6)	Colorectal cancer (0.76)	HIV (27.64)	CKD (2.69)
Florida	IHD (0.8)	Lung cancer (1.28)	COPD (2.11)	Stroke (0.62)	Alzheimer (1.39)	Road injury (1.3)	Self-harm (0.99)	Drugs (7.69)	Colorectal cancer (0.9)	Diabetes (2.3)
Georgia	IHD (0.75)	Lung cancer (1.16)	Stroke (0.75)	COPD (1.98)	Road injury (1.4)	Self-harm (0.94)	LRI (1.61)	Colorectal cancer (0.9)	CKD (2.77)	Alzheimer (0.9)
Hawaii	IHD (0.6)	Lung cancer (0.88)	Alzheimer (1.47)	Stroke (0.6)	Self-harm (0.98)	Colorectal cancer (0.84)	Road injury (0.86)	LRI (1.21)	COPD (0.94)	CKD (1.93)
Idaho	IHD (0.59)	Lung cancer (0.97)	COPD (2.01)	Self-harm (1.23)	Stroke (0.55)	Road injury (1.19)	Alzheimer (1.14)	Diabetes (1.73)	Colorectal cancer (0.75)	LRI (0.88)
Illinois	IHD (0.75)	Lung cancer (1.09)	Stroke (0.59)	COPD (1.68)	Alzheimer (1.09)	Self-harm (0.81)	Road injury (0.95)	Colorectal cancer (0.83)	LRI (1.41)	Drugs (6.97)
Indiana	IHD (0.87)	Lung cancer (1.44)	COPD (2.5)	Stroke (0.67)	Self-harm (1.15)	Alzheimer (1.34)	Road injury (1.25)	Drugs (7.61)	Colorectal cancer (0.94)	Diabetes (2.29)
Iowa	IHD (0.79)	Lung cancer (1.21)	COPD (2.2)	Stroke (0.62)	Alzheimer (1.16)	Road injury (1.13)	Self-harm (0.94)	Colorectal cancer (0.91)	LRI (1.19)	Diabetes (1.73)
Kansas	IHD (0.73)	Lung cancer (1.18)	COPD (2.31)	Stroke (0.65)	Road injury (1.39)	Self-harm (1.1)	Alzheimer (1.18)	Colorectal cancer (0.83)	LRI (1.33)	Diabetes (1.93)
Kentucky	IHD (1.04)	Lung cancer (1.97)	COPD (3.04)	Stroke (0.72)	Road injury (1.54)	Drugs (12.44)	Self-harm (1.18)	Alzheimer (1.39)	Colorectal cancer (1.18)	LRI (1.65)
Louisiana	IHD (0.98)	Lung cancer (1.46)	Stroke (0.84)	Road injury (1.73)	COPD (2.03)	Self-harm (1.07)	CKD (3.73)	LRI (1.81)	Violence (5.13)	Colorectal cancer (1.12)
Maine	IHD (0.75)	Lung cancer (1.59)	COPD (2.75)	Stroke (0.66)	Alzheimer (1.48)	Self-harm (1.07)	Road injury (1.15)	Colorectal cancer (0.98)	Diabetes (2.39)	LRI (1.3)
Maryland	IHD (0.72)	Lung cancer (0.98)	Self-harm (1.26)	Stroke (0.57)	Alzheimer (1.16)	COPD (1.39)	Road injury (1.05)	LRI (1.46)	Colorectal cancer (0.78)	Violence (4.2)
Massachusetts	IHD (0.64)	Lung cancer (1.05)	Alzheimer (1.27)	Stroke (0.52)	COPD (1.64)	Drugs (9.07)	Self-harm (0.89)	LRI (1.53)	Colorectal cancer (0.76)	Road injury (0.82)
Michigan	IHD (0.96)	Lung cancer (1.31)	COPD (2.1)	Stroke (0.62)	Alzheimer (1.28)	Self-harm (1.06)	Drugs (8.98)	Road injury (1.06)	Colorectal cancer (0.92)	Diabetes (2.25)
Minnesota	IHD (0.49)	Lung cancer (0.97)	Stroke (0.54)	COPD (1.62)	Alzheimer (1.14)	Self-harm (0.93)	Road injury (0.93)	Colorectal cancer (0.72)	Diabetes (1.67)	CKD (1.83)
Mississippi	IHD (1.14)	Lung cancer (1.75)	Road injury (2.19)	Stroke (0.94)	COPD (2.54)	LRI (1.89)	CKD (3.79)	Self-harm (1.05)	Colorectal cancer (1.19)	Diabetes (2.75)
Missouri	IHD (0.97)	Lung cancer (1.47)	COPD (2.49)	Stroke (0.7)	Road injury (1.43)	Self-harm (1.14)	Alzheimer (1.32)	Drugs (8.28)	Colorectal cancer (0.93)	LRI (1.44)
Montana	IHD (0.69)	Lung cancer (1.14)	COPD (2.53)	Self-harm (1.46)	Road injury (1.71)	Stroke (0.6)	Alzheimer (1.38)	Colorectal cancer (0.86)	Diabetes (1.94)	LRI (1.12)
Nebraska	IHD (0.6)	Lung cancer (1.1)	COPD (2.15)	Stroke (0.6)	Road injury (1.28)	Alzheimer (1.16)	Self-harm (0.94)	Colorectal cancer (0.86)	Diabetes (1.77)	LRI (1.04)
Nevada	IHD (0.75)	Lung cancer (1.16)	COPD (2.22)	Self-harm (1.33)	Stroke (0.59)	Drugs (9.7)	Road injury (1.04)	LRI (1.49)	Colorectal cancer (0.93)	Alzheimer (0.89)
New Hampshire	IHD (0.68)	Lung cancer (1.23)	COPD (2.29)	Alzheimer (1.36)	Self-harm (1.17)	Stroke (0.52)	Road injury (1.1)	Colorectal cancer (0.82)	Drug (7.22)	Diabetes (2.15)
New Jersey	IHD (0.72)	Lung cancer (0.93)	Stroke (0.51)	Alzheimer (1.07)	COPD (1.37)	Drugs (7.52)	Colorectal cancer (0.84)	Diabetes (2.21)	LRI (1.32)	Self-harm (0.69)
New Mexico	IHD (0.68)	Road injury (1.69)	Self-harm (1.46)	Lung cancer (0.85)	Drugs (12.78)	COPD (2.03)	Alzheimer (1.3)	Stroke (0.51)	Diabetes (2.65)	Colorectal cancer (0.91)
New York	IHD (0.82)	Lung cancer (0.92)	Alzheimer (1.12)	COPD (1.34)	Stroke (0.38)	Colorectal cancer (0.73)	Drugs (6.23)	Self-harm (0.69)	LRI (1.29)	Road injury (0.76)
North Carolina	IHD (0.76)	Lung cancer (1.32)	Stroke (0.76)	COPD (2.12)	Road injury (1.35)	Self-harm (0.94)	Alzheimer (1.12)	LRI (1.52)	CKD (2.71)	Drugs (6.97)
North Dakota	IHD (0.64)	Lung cancer (0.94)	Road injury (1.55)	Self-harm (1.11)	Alzheimer (1.25)	Stroke (0.52)	COPD (1.47)	Colorectal cancer (0.75)	Diabetes (1.87)	Congenital (1.5)
Ohio	IHD (0.9)	Lung cancer (1.44)	COPD (2.47)	Stroke (0.71)	Alzheimer (1.35)	Drugs (10.5)	Self-harm (1.01)	Road injury (1.08)	Diabetes (2.66)	Colorectal cancer (0.98)
Oklahoma	IHD (1.1)	Lung cancer (1.48)	COPD (2.9)	Road injury (1.8)	Stroke (0.76)	Self-harm (1.31)	Alzheimer (1.29)	Drugs (9.48)	Diabetes (2.5)	LRI (1.5)
Oregon	IHD (0.57)	Lung cancer (1.14)	COPD (2.21)	Stroke (0.67)	Self-harm (1.23)	Alzheimer (1.2)	Road injury (0.99)	Diabetes (2.27)	Colorectal cancer (0.81)	Drugs (6.59)
Pennsylvania	IHD (0.87)	Lung cancer (1.27)	Stroke (0.67)	COPD (2.02)	Alzheimer (1.36)	Drugs (10.76)	Self-harm (1.03)	Road injury (1.13)	Colorectal cancer (0.98)	LRI (1.54)
Rhode Island	IHD (0.8)	Lung cancer (1.23)	Alzheimer (1.45)	COPD (1.87)	Stroke (0.5)	Drugs (8.34)	Self-harm (0.91)	Colorectal cancer (0.84)	Diabetes (1.93)	LRI (1.17)
South Carolina	IHD (0.9)	Lung cancer (1.48)	Stroke (0.92)	COPD (2.45)	Road injury (1.71)	Self-harm (1.0)	Alzheimer (1.2)	LRI (1.61)	CKD (3.07)	Colorectal cancer (1.01)
South Dakota	IHD (0.73)	Lung cancer (1.12)	Road injury (1.52)	COPD (1.89)	Stroke (0.57)	Self-harm (1.07)	Alzheimer (1.17)	Colorectal cancer (0.88)	Diabetes (1.88)	LRI (1.1)
Tennessee	IHD (1.04)	Lung cancer (1.66)	COPD (2.6)	Stroke (0.85)	Road injury (1.6)	Self-harm (1.16)	Drugs (9.48)	Alzheimer (1.2)	LRI (1.7)	Colorectal cancer (1.03)
Texas	IHD (0.65)	Lung cancer (0.87)	Road injury (1.34)	Stroke (0.56)	COPD (1.47)	Self-harm (0.89)	Alzheimer (1.01)	LRI (1.15)	Colorectal cancer (0.74)	CKD (2.18)
Utah	IHD (0.4)	Self-harm (1.58)	Stroke (0.43)	Road injury (0.87)	COPD (1.08)	Lung cancer (0.43)	Drugs (6.29)	Diabetes (1.62)	Alzheimer (0.71)	LRI (0.89)
Vermont	IHD (0.68)	Lung cancer (1.26)	COPD (2.4)	Alzheimer (1.32)	Self-harm (1.15)	Stroke (0.54)	Road injury (1.09)	Colorectal cancer (0.84)	Diabetes (2.09)	Drugs (5.22)
Virginia	IHD (0.68)	Lung cancer (1.07)	Stroke (0.63)	COPD (1.69)	Alzheimer (1.11)	Self-harm (0.98)	Road injury (1.08)	LRI (1.5)	Colorectal cancer (0.79)	CKD (2.52)
Washington	IHD (0.55)	Lung cancer (0.93)	COPD (1.77)	Stroke (0.55)	Self-harm (1.08)	Alzheimer (0.99)	Drugs (7.52)	Road injury (0.89)	Colorectal cancer (0.7)	Diabetes (1.91)
West Virginia	IHD (1.16)	Lung cancer (2.05)	COPD (3.53)	Stroke (0.83)	Drugs (14.38)	Road injury (1.6)	Diabetes (3.67)	Self-harm (1.26)	Alzheimer (1.49)	Colorectal cancer (1.33)
Wisconsin	IHD (0.7)	Lung cancer (1.12)	Stroke (0.61)	COPD (1.79)	Alzheimer (1.22)	Self-harm (1.03)	Road injury (1.09)	Colorectal cancer (0.8)	Drugs (6.27)	LRI (1.1)
Wyoming	IHD (0.67)	COPD (2.6)	Lung cancer (0.93)	Road injury (1.84)	Self-harm (1.5)	Stroke (0.51)	Alzheimer (0.97)	Drugs (6.69)	Colorectal cancer (0.79)	LRI (1.26)

Ratio of observed to expected YLLs
(0.0-0.66)
(0.67-0.83)
(0.84-0.95)
(0.96-1.10)
(1.11-1.22)
(1.23-1.44)
(1.45-1.81)
(1.82-2.56)
>2.56

Ratio details: Alabama's, stroke ratio (eg, 1.0 [observed and expected rates were similar]; diabetes [2.85 × above expected]). See Appendix Table 2 in Supplement 2 for explanation of terms.

Figure 8. Ratio of Observed Years Lived With Disability (YLDs) to Expected YLDs Based on the Sociodemographic Index (SDI) for the United States Overall, the 50 States, and the District of Columbia in 2016 for the 10 Leading Causes in Each Jurisdiction

	1	2	3	4	5	6	7	8	9	10
United States	Back and neck (0.86)	Skin (1.19)	Depression (1.29)	Sense organ (0.85)	Drugs (3.49)	Diabetes (1.52)	Other MSK (2.37)	Migraine (0.85)	Anxiety (1.25)	Falls (0.85)
Alabama	Back and neck (1.09)	Skin (1.07)	Depression (1.29)	Diabetes (1.9)	Sense organ (0.88)	Drugs (3.37)	Other MSK (2.39)	Migraine (0.89)	Anxiety (1.27)	COPD (3.72)
Alaska	Back and neck (0.86)	Depression (1.36)	Skin (1.03)	Drugs (3.43)	Other MSK (2.1)	Sense organ (0.65)	Migraine (0.76)	Anxiety (1.24)	Diabetes (1.16)	Falls (0.79)
Arizona	Back and neck (0.99)	Skin (1.2)	Depression (1.44)	Drugs (4.39)	Sense organ (0.9)	Migraine (0.92)	Other MSK (2.42)	Diabetes (1.51)	Anxiety (1.25)	Falls (0.94)
Arkansas	Back and neck (0.84)	Depression (1.47)	Skin (0.99)	Sense organ (0.88)	Drugs (3.53)	Diabetes (1.6)	Migraine (0.83)	Anxiety (1.26)	Other MSK (1.97)	Falls (0.94)
California	Back and neck (0.74)	Skin (1.25)	Depression (1.23)	Sense organ (0.78)	Migraine (0.85)	Drugs (2.97)	Other MSK (2.19)	Diabetes (1.24)	Anxiety (1.2)	Falls (0.77)
Colorado	Back and neck (0.8)	Depression (1.48)	Skin (1.1)	Drugs (3.38)	Sense organ (0.79)	Other MSK (2.45)	Migraine (0.88)	Anxiety (1.28)	Falls (0.9)	Diabetes (0.87)
Connecticut	Back and neck (0.85)	Skin (1.34)	Drugs (4.23)	Sense organ (0.89)	Depression (1.17)	Diabetes (1.52)	Other MSK (2.42)	Migraine (0.87)	Anxiety (1.28)	Falls (0.86)
Delaware	Back and neck (0.88)	Skin (1.11)	Drugs (4.36)	Depression (1.32)	Diabetes (1.92)	Sense organ (0.89)	Other MSK (2.42)	Migraine (0.87)	Anxiety (1.27)	Falls (0.85)
District of Columbia	Skin (1.43)	Back and neck (0.57)	Violence (4.39)	Depression (1.16)	Other MSK (2.24)	Sense organ (0.69)	Migraine (0.8)	Anxiety (1.32)	Diabetes (0.94)	Falls (0.71)
Florida	Back and neck (1.03)	Skin (1.22)	Sense organ (1.01)	Depression (1.27)	Diabetes (1.83)	Drugs (3.48)	Migraine (0.84)	Other MSK (2.25)	Anxiety (1.27)	Falls (0.97)
Georgia	Back and neck (0.86)	Skin (1.23)	Depression (1.21)	Diabetes (1.6)	Sense organ (0.77)	Drugs (3.07)	Migraine (0.87)	Other MSK (2.24)	Anxiety (1.28)	Falls (0.81)
Hawaii	Back and neck (0.73)	Skin (1.28)	Depression (1.29)	Sense organ (0.91)	Diabetes (1.52)	Migraine (0.81)	Anxiety (1.25)	Other MSK (1.98)	Drugs (2.5)	Falls (0.86)
Idaho	Back and neck (0.87)	Depression (1.45)	Skin (1.08)	Sense organ (0.85)	Drugs (3.27)	Other MSK (2.39)	Migraine (0.82)	Diabetes (1.32)	Anxiety (1.24)	Falls (0.97)
Illinois	Back and neck (0.85)	Skin (1.11)	Depression (1.16)	Sense organ (0.82)	Drugs (3.24)	Other MSK (2.4)	Diabetes (1.31)	Migraine (0.79)	Anxiety (1.28)	Falls (0.77)
Indiana	Back and neck (0.96)	Skin (1.1)	Depression (1.39)	Drugs (3.51)	Diabetes (1.67)	Sense organ (0.84)	Other MSK (2.42)	Migraine (0.89)	Anxiety (1.26)	COPD (3.86)
Iowa	Back and neck (0.93)	Skin (1.07)	Depression (1.23)	Sense organ (0.9)	Other MSK (2.54)	Diabetes (1.36)	Migraine (0.8)	Anxiety (1.26)	Falls (0.94)	Drugs (1.98)
Kansas	Back and neck (0.85)	Skin (1.07)	Depression (1.32)	Sense organ (0.83)	Other MSK (2.32)	Diabetes (1.44)	Migraine (0.78)	Anxiety (1.25)	Drugs (2.4)	Falls (0.86)
Kentucky	Back and neck (1.08)	Drugs (5.69)	Skin (1.13)	Depression (1.39)	Diabetes (1.81)	Sense organ (0.88)	Other MSK (2.45)	Migraine (0.91)	COPD (4.73)	Anxiety (1.26)
Louisiana	Back and neck (0.97)	Skin (1.18)	Drugs (4.33)	Depression (1.23)	Diabetes (1.64)	Sense organ (0.81)	Other MSK (2.22)	Migraine (0.81)	Anxiety (1.26)	Falls (0.76)
Maine	Back and neck (0.83)	Depression (1.41)	Sense organ (1.02)	Skin (1.01)	Other MSK (2.83)	Diabetes (1.6)	Drugs (3.28)	Migraine (0.81)	Anxiety (1.28)	Falls (1.02)
Maryland	Back and neck (0.83)	Skin (1.26)	Depression (1.32)	Sense organ (0.81)	Diabetes (1.55)	Other MSK (2.42)	Migraine (0.86)	Drugs (2.97)	Anxiety (1.28)	Falls (0.75)
Massachusetts	Back and neck (0.72)	Skin (1.29)	Drugs (4.56)	Depression (1.28)	Sense organ (0.86)	Other MSK (2.55)	Migraine (0.88)	Diabetes (1.27)	Anxiety (1.29)	Falls (0.81)
Michigan	Back and neck (0.77)	Skin (1.18)	Depression (1.33)	Diabetes (1.78)	Sense organ (0.88)	Drugs (3.55)	Other MSK (2.64)	Migraine (0.83)	Anxiety (1.27)	COPD (3.85)
Minnesota	Back and neck (0.63)	Skin (1.08)	Depression (1.27)	Sense organ (0.83)	Other MSK (2.24)	Migraine (0.79)	Anxiety (1.27)	Diabetes (1.21)	Falls (0.93)	Drugs (2.32)
Mississippi	Back and neck (0.8)	Skin (0.98)	Depression (1.25)	Diabetes (1.77)	Sense organ (0.85)	Drugs (3.05)	Migraine (0.8)	Anxiety (1.28)	Other MSK (1.9)	Falls (0.92)
Missouri	Back and neck (0.88)	Skin (1.08)	Depression (1.37)	Sense organ (0.89)	Diabetes (1.63)	Other MSK (2.44)	Drugs (3.07)	Migraine (0.85)	Anxiety (1.26)	Falls (0.93)
Montana	Back and neck (0.88)	Depression (1.46)	Skin (1.03)	Sense organ (0.94)	Other MSK (2.48)	Migraine (0.79)	Drugs (2.63)	Anxiety (1.25)	Falls (1.07)	Diabetes (1.23)
Nebraska	Back and neck (0.86)	Skin (1.07)	Depression (1.21)	Sense organ (0.83)	Other MSK (2.21)	Diabetes (1.38)	Migraine (0.75)	Anxiety (1.25)	Falls (0.91)	Drugs 1.88)
Nevada	Back and neck (0.99)	Depression (1.44)	Skin (1.08)	Drugs (4.13)	Sense organ (0.84)	Diabetes (1.55)	Migraine (0.86)	Other MSK (2.17)	Anxiety (1.27)	COPD (3.45)
New Hampshire	Back and neck (0.84)	Drugs (4.57)	Depression (1.44)	Skin (1.08)	Sense organ (0.91)	Other MSK (2.6)	Diabetes (1.43)	Migraine (0.86)	Anxiety (1.29)	Falls (0.95)
New Jersey	Back and neck (0.84)	Skin (1.3)	Sense organ (0.87)	Drugs (3.55)	Depression (1.1)	Diabetes (1.58)	Other MSK (2.42)	Migraine (0.77)	Anxiety (1.28)	Falls (0.74)
New Mexico	Back and neck (0.92)	Depression (1.47)	Drugs (4.35)	Skin (1.06)	Sense organ (0.92)	Diabetes (1.62)	Other MSK (2.44)	Migraine (0.9)	Falls (1.12)	Anxiety (1.25)
New York	Back and neck (0.87)	Skin (1.4)	Depression (1.47)	Diabetes (1.72)	Drugs (3.69)	Sense organ (0.87)	Other MSK (2.64)	Migraine (0.86)	Anxiety (1.28)	COPD (3.47)
North Carolina	Back and neck (0.79)	Skin (1.13)	Depression (1.18)	Sense organ (0.85)	Diabetes (1.65)	Drugs (3.33)	Other MSK (2.28)	Migraine (0.86)	Anxiety (1.08)	Falls (0.9)
North Dakota	Back and neck (0.75)	Skin (1.02)	Other MSK (3.07)	Depression (1.2)	Sense organ (0.8)	Migraine (0.81)	Diabetes (1.29)	Anxiety (1.24)	Drugs (2.33)	Falls (0.9)
Ohio	Back and neck (0.92)	Skin (1.14)	Drugs (4.41)	Depression (1.32)	Diabetes (1.75)	Sense organ (0.89)	Other MSK (2.44)	Migraine (0.9)	Anxiety (1.26)	COPD (4.1)
Oklahoma	Back and neck (1.0)	Skin (1.12)	Depression (1.39)	Drugs (4.35)	Sense organ (0.84)	Diabetes (1.58)	Other MSK (2.38)	Migraine (0.87)	Anxiety (1.25)	Falls (0.96)
Oregon	Back and neck (0.81)	Skin (1.21)	Drugs (4.38)	Depression (1.37)	Sense organ (0.89)	Other MSK (2.61)	Migraine (0.93)	Diabetes (1.3)	Falls (0.88)	Anxiety (1.01)
Pennsylvania	Back and neck (1.01)	Skin (1.18)	Drugs (4.47)	Depression (1.26)	Sense organ (0.93)	Diabetes (1.66)	Other MSK (2.63)	Migraine (0.91)	Anxiety (1.26)	Falls (0.95)
Rhode Island	Back and neck (0.77)	Drugs (5.45)	Skin (1.22)	Depression (1.4)	Sense organ (0.93)	Diabetes (1.52)	Migraine (0.92)	Other MSK (2.35)	Anxiety (1.29)	Falls (0.96)
South Carolina	Back and neck (0.94)	Skin (1.21)	Depression (1.31)	Drugs (3.87)	Diabetes (1.85)	Sense organ (0.89)	Other MSK (2.4)	Migraine (0.87)	Anxiety (1.27)	Falls (0.9)
South Dakota	Back and neck (0.86)	Skin (1.07)	Sense organ (0.88)	Depression (1.14)	Other MSK (2.62)	Diabetes (1.31)	Anxiety (1.24)	Migraine (0.74)	Falls (1.07)	Drugs (2.15)
Tennessee	Back and neck (0.9)	Skin (1.07)	Depression (1.39)	Drugs (4.37)	Diabetes (1.73)	Sense organ (0.86)	Other MSK (2.34)	Migraine (0.86)	Anxiety (1.27)	Falls (0.87)
Texas	Back and neck (0.87)	Skin (1.23)	Depression (1.13)	Sense organ (0.74)	Migraine (0.85)	Diabetes (1.44)	Other MSK (2.17)	Drugs (2.8)	Anxiety (1.27)	Falls (0.77)
Utah	Back and neck (0.8)	Depression (1.62)	Skin (1.15)	Drugs (3.79)	Migraine (0.86)	Other MSK (2.21)	Sense organ (0.68)	Anxiety (1.23)	Diabetes (1.15)	Falls (0.75)
Vermont	Back and neck (0.76)	Depression (1.37)	Skin (1.02)	Sense (0.95)	Drugs (3.48)	Other MSK (2.43)	Migraine (0.84)	Diabetes (1.33)	Falls (1.06)	Anxiety (1.29)
Virginia	Back and neck (0.83)	Skin (1.21)	Depression (1.38)	Sense (0.81)	Diabetes (1.55)	Migraine (0.92)	Other MSK (2.39)	Drugs (2.88)	Anxiety (1.28)	Falls (0.78)
Washington	Back and neck (0.78)	Skin (1.18)	Depression (1.44)	Drugs (3.58)	Other MSK (2.62)	Sense organ (0.81)	Migraine (0.89)	Diabetes (1.27)	Anxiety (1.27)	Falls (0.87)
West Virginia	Back and neck (1.04)	Drugs (5.68)	Depression (2.23)	Depression (1.42)	Skin (1.08)	Sense organ (0.98)	COPD (5.53)	Migraine (0.92)	Other MSK (2.38)	Anxiety (1.26)
Wisconsin	Back and neck (0.85)	Skin (1.08)	Depression (1.36)	Sense organ (0.88)	Drugs (3.4)	Other MSK (2.32)	Migraine (0.84)	Diabetes (1.38)	Anxiety (1.27)	Falls (1.0)
Wyoming	Back and neck (1.0)	Depression (1.47)	Skin (0.97)	Sense organ (0.8)	Other MSK (2.29)	Drugs (2.91)	Migraine (0.79)	Anxiety (1.24)	Diabetes (1.12)	Falls (0.95)

Ratio of observed to expected YLDs
(0.0-0.66)
(0.67-0.83)
(0.84-0.95)
(0.96-1.10)
(1.11-1.22)
(1.23-1.44)
(1.45-1.81)
(1.82-2.56)
>2.56

See Figure 7 caption for details. See Appendix Table 2 in Supplement 2 for explanation of terms.

Figure 9. Ratio of Observed Disability-Adjusted Life-Years (DALYs) to Expected DALYs Based on the Sociodemographic Index (SDI) for the United States Overall, the 50 States, and the District of Columbia in 2016 for the 10 Leading Causes in Each Jurisdiction

	1	2	3	4	5	6	7	8	9	10
United States	IHD (0.74)	Back and neck (0.86)	Drugs (4.37)	Lung cancer (1.11)	COPD (2.16)	Diabetes (1.68)	Skin (1.21)	Stroke (0.66)	Depression (1.29)	Road injury (1.11)
Alabama	IHD (1.02)	Back and neck (1.09)	Lung cancer (1.72)	COPD (3.09)	Stroke (1.02)	Diabetes (2.21)	Road injury (1.75)	Drugs (4.37)	Skin (1.11)	Depression (1.29)
Alaska	IHD (0.49)	Back and neck (0.86)	Drugs (4.64)	Lung cancer (1.0)	Self-harm (1.63)	Skin (1.04)	Depression (1.36)	Diabetes (1.31)	COPD (1.56)	Stroke (0.5)
Arizona	IHD (0.64)	Back and neck (0.99)	Drugs (5.4)	COPD (2.13)	Diabetes (1.66)	Skin (1.21)	Lung cancer (0.95)	Depression (1.44)	Stroke (0.58)	Road injury (1.18)
Arkansas	IHD (1.1)	Lung cancer (1.77)	COPD (2.9)	Back and neck (0.84)	Stroke (0.96)	Diabetes (1.92)	Road injury (1.64)	Drugs (3.87)	Depression (1.47)	Skin (1.02)
California	IHD (0.57)	Back and neck (0.74)	Skin (1.26)	Drugs (3.69)	Diabetes (1.33)	Depression (1.23)	Stroke (0.54)	COPD (1.49)	Lung cancer (0.68)	Sense organ (0.78)
Colorado	IHD (0.48)	Back and neck (0.8)	Drugs (4.29)	COPD (2.06)	Depression (1.48)	Skin (1.12)	Self-harm (1.41)	Stroke (0.51)	Road injury (1.08)	Other MSK (2.39)
Connecticut	IHD (0.63)	Back and neck (0.85)	Drugs (5.0)	Skin (1.36)	Lung cancer (0.98)	Diabetes (1.58)	COPD (1.99)	Alzheimer (1.29)	Stroke (0.57)	Sense organ (0.89)
Delaware	IHD (0.78)	Back and neck (0.88)	Drugs (5.25)	Lung cancer (1.35)	Diabetes (2.05)	COPD (2.34)	Stroke (0.67)	Skin (1.14)	Depression (1.32)	Road injury (1.19)
District of Columbia	IHD (0.72)	Drugs (5.16)	Skin (1.48)	Back and neck (0.57)	Lung cancer (0.85)	Diabetes (1.28)	Violence (6.16)	Depression (1.16)	Stroke (0.49)	Alzheimer (0.92)
Florida	IHD (0.81)	Back and neck (1.03)	Lung cancer (1.28)	COPD (2.57)	Diabetes (1.98)	Drugs (4.48)	Skin (1.24)	Stroke (0.71)	Alzheimer (1.4)	Sense organ (1.01)
Georgia	IHD (0.73)	Back and neck (0.86)	Lung cancer (1.17)	Stroke (0.77)	COPD (2.19)	Diabetes (1.74)	Skin (1.26)	Drugs (3.84)	Road injury (1.26)	Depression (1.21)
Hawaii	IHD (0.59)	Back and neck (0.73)	Skin (1.3)	Alzheimer (1.5)	Stroke (0.67)	Diabetes (1.54)	Lung cancer (0.88)	Depression (1.29)	Drugs (3.01)	Sense organ (0.91)
Idaho	IHD (0.58)	Back and neck (0.87)	COPD (2.12)	Drugs (3.61)	Diabetes (1.46)	Depression (1.45)	Lung cancer (0.97)	Skin (1.1)	Stroke (0.62)	Road injury (1.14)
Illinois	IHD (0.73)	Back and neck (0.85)	Lung cancer (1.1)	Drugs (4.12)	Skin (1.13)	COPD (1.93)	Stroke (0.65)	Diabetes (1.48)	Depression (1.16)	Alzheimer (1.09)
Indiana	IHD (0.86)	Back and neck (0.96)	COPD (2.81)	Lung cancer (1.44)	Drugs (4.49)	Diabetes (1.87)	Stroke (0.73)	Skin (1.12)	Depression (1.39)	Alzheimer (1.36)
Iowa	IHD (0.77)	Back and neck (0.93)	Lung cancer (1.21)	COPD (2.36)	Stroke (0.69)	Diabetes (1.48)	Skin (1.09)	Depression (1.23)	Alzheimer (1.17)	Sense organ (0.9)
Kansas	IHD (0.72)	Back and neck (0.85)	COPD (2.43)	Lung cancer (1.19)	Stroke (0.7)	Diabetes (1.59)	Skin (1.09)	Road injury (1.29)	Depression (1.32)	Drugs (3.02)
Kentucky	IHD (1.02)	Drugs (7.31)	Lung cancer (1.98)	Back and neck (1.08)	COPD (3.43)	Diabetes (2.01)	Stroke (0.78)	Road injury (1.42)	Skin (1.15)	Depression (1.39)
Louisiana	IHD (0.96)	Back and neck (0.97)	Lung cancer (1.47)	Drugs (5.36)	Diabetes (2.0)	Stroke (0.86)	COPD (2.24)	Road injury (1.54)	Skin (1.22)	CKD (3.05)
Maine	IHD (0.74)	Lung cancer (1.6)	Back and neck (0.83)	COPD (2.95)	Diabetes (1.85)	Stroke (0.74)	Drugs (3.96)	Alzheimer (1.47)	Depression (1.41)	Skin (1.03)
Maryland	IHD (0.7)	Back and neck (0.83)	Skin (1.29)	Diabetes (1.71)	Lung cancer (0.98)	Stroke (0.62)	Depression (1.32)	COPD (1.77)	Drugs (3.08)	Alzheimer (1.16)
Massachusetts	IHD (0.63)	Drugs (5.61)	Back and neck (0.72)	Lung cancer (1.06)	Skin (1.31)	COPD (2.0)	Alzheimer (1.24)	Diabetes (1.37)	Stroke (0.59)	Depression (1.28)
Michigan	IHD (0.95)	Back and neck (0.77)	Lung cancer (1.31)	Drugs (4.84)	COPD (2.51)	Diabetes (1.93)	Stroke (0.69)	Skin (1.2)	Depression (1.33)	Alzheimer (1.29)
Minnesota	IHD (0.48)	Back and neck (0.63)	Lung cancer (0.97)	Skin (1.09)	Stroke (0.61)	Diabetes (1.35)	COPD (1.74)	Depression (1.27)	Alzheimer (1.13)	IHD (2.73)
Mississippi	IHD (1.11)	Lung cancer (1.75)	Stroke (0.96)	Road injury (1.93)	COPD (2.63)	Back and neck (0.8)	Diabetes (2.1)	Drugs (3.78)	CKD (3.14)	Skin (1.03)
Missouri	IHD (0.94)	Back and neck (0.88)	Lung cancer (1.48)	COPD (2.73)	Drugs (4.31)	Diabetes (1.78)	Stroke (0.75)	Skin (1.11)	Road injury (1.31)	Depression (1.37)
Montana	IHD (0.68)	Back and neck (0.88)	COPD (2.69)	Lung cancer (1.15)	Road injury (1.56)	Stroke (0.67)	Depression (1.46)	Diabetes (1.45)	Alzheimer (1.38)	Self-harm (1.48)
Nebraska	IHD (0.59)	Back and neck (0.86)	COPD (2.3)	Lung cancer (1.1)	Stroke (0.66)	Diabetes (1.51)	Skin (1.08)	Road injury (1.2)	Depression (1.21)	Alzheimer (1.16)
Nevada	IHD (0.73)	Back and neck (0.99)	Drugs (5.46)	COPD (2.5)	Lung cancer (1.16)	Stroke (0.65)	Diabetes (1.49)	Depression (1.44)	Skin (1.1)	Self-harm (1.34)
New Hampshire	IHD (0.67)	Back and neck (0.84)	Drugs (5.19)	Lung cancer (1.24)	COPD (2.55)	Diabetes (1.64)	Alzheimer (1.34)	Depression (1.44)	Skin (1.1)	Stroke (0.6)
New Jersey	IHD (0.71)	Back and neck (0.84)	Drugs (4.48)	Skin (1.33)	Diabetes (1.77)	Lung cancer (0.93)	COPD (1.77)	Stroke (0.57)	Sense organ (0.87)	Alzheimer (1.07)
New Mexico	IHD (0.67)	Drugs (6.37)	Back and neck (0.92)	Diabetes (1.96)	COPD (2.26)	Road injury (1.53)	Depression (1.47)	Skin (1.08)	Self-harm (1.48)	Stroke (0.58)
New York	IHD (0.81)	Back and neck (0.87)	Skin (1.42)	Drugs (4.28)	Diabetes (1.74)	Lung cancer (0.93)	Depression (1.47)	COPD (1.86)	Alzheimer (1.13)	Sense organ (0.87)
North Carolina	IHD (0.74)	Back and neck (0.79)	Lung cancer (1.32)	Stroke (0.83)	COPD (2.26)	Diabetes (1.81)	Drugs (4.2)	Skin (1.16)	Road injury (1.27)	Depression (1.18)
North Dakota	IHD (0.62)	Back and neck (0.75)	Road injury (1.42)	Lung cancer (0.95)	Diabetes (1.47)	Other MSK (2.97)	Stroke (0.59)	Skin (1.03)	COPD (1.7)	Alzheimer (1.27)
Ohio	IHD (0.88)	Back and neck (0.92)	Drugs (5.86)	Lung cancer (1.45)	COPD (2.85)	Diabetes (2.04)	Stroke (0.76)	Skin (1.17)	Alzheimer (1.35)	Depression (1.32)
Oklahoma	IHD (1.08)	Back and neck (1.0)	COPD (3.1)	Drugs (5.58)	Lung cancer (1.49)	Diabetes (1.88)	Stroke (0.81)	Road injury (1.61)	Skin (1.15)	Depression (1.39)
Oregon	IHD (0.56)	Back and neck (0.81)	Drugs (4.9)	COPD (2.39)	Lung cancer (1.14)	Stroke (0.73)	Skin (1.23)	Diabetes (1.6)	Depression (1.37)	Alzheimer (1.19)
Pennsylvania	IHD (0.85)	Back and neck (1.01)	Drugs (5.95)	Lung cancer (1.28)	COPD (2.41)	Diabetes (1.86)	Stroke (0.73)	Skin (1.21)	Alzheimer (1.36)	Depression (1.26)
Rhode Island	IHD (0.78)	Drugs (6.13)	Back and neck (0.77)	Lung cancer (1.24)	COPD (2.23)	Skin (1.24)	Diabetes (1.64)	Alzheimer (1.43)	Depression (1.4)	Stroke (0.58)
South Carolina	IHD (0.88)	Back and neck (0.94)	Lung cancer (1.49)	Stroke (0.94)	COPD (2.65)	Diabetes (2.03)	Drugs (4.64)	Road injury (1.54)	Skin (1.24)	Depression (1.31)
South Dakota	IHD (0.72)	Back and neck (0.86)	Lung cancer (1.12)	COPD (2.07)	Road injury (1.42)	Stroke (0.65)	Diabetes (1.5)	Skin (1.08)	Other MSK (2.61)	Alzheimer (1.18)
Tennessee	IHD (1.02)	Lung cancer (1.67)	Back and neck (0.9)	Drugs (5.59)	COPD (2.81)	Stroke (0.89)	Diabetes (1.96)	Road injury (1.45)	Skin (1.12)	Depression (1.39)
Texas	IHD (0.65)	Back and neck (0.87)	Skin (1.25)	Diabetes (1.54)	Drugs (3.46)	COPD (1.72)	Stroke (0.6)	Road injury (1.24)	Lung cancer (0.88)	Depression (1.13)
Utah	Back and neck (0.8)	IHD (0.4)	Drugs (4.39)	Depression (1.62)	Self-harm (1.59)	Skin (1.16)	Diabetes (1.31)	Stroke (0.48)	Other MSK (2.21)	COPD (1.17)
Vermont	IHD (0.67)	Back and neck (0.76)	Lung cancer (1.27)	COPD (2.58)	Drugs (3.88)	Diabetes (1.55)	Alzheimer (1.31)	Stroke (0.61)	Depression (1.37)	Skin (1.04)
Virginia	IHD (0.66)	Back and neck (0.83)	Lung cancer (1.08)	Diabetes (1.67)	Skin (1.23)	Stroke (0.67)	COPD (1.94)	Drugs (3.46)	Depression (1.38)	Alzheimer (1.11)
Washington	IHD (0.53)	Back and neck (0.78)	Drugs (4.51)	Skin (1.2)	Lung cancer (0.93)	Depression (1.44)	Diabetes (1.47)	COPD (1.88)	Stroke (0.61)	Other MSK (2.53)
West Virginia	IHD (1.14)	COPD (3.99)	Drugs (7.77)	Lung cancer (205)	Back and neck (1.04)	Diabetes (2.72)	Stroke (0.89)	Road injury (1.48)	Alzheimer (1.5)	Skin (1.11)
Wisconsin	IHD (0.69)	Back and neck (0.85)	Lung cancer (1.13)	Drugs (4.08)	COPD (1.96)	Stroke (0.67)	Diabetes (1.48)	Skin (1.1)	Depression (1.36)	Alzheimer (1.22)
Wyoming	IHD (0.66)	Back and neck (1.0)	Lung cancer (2.78)	Road injury (1.63)	Drugs (3.8)	Lung cancer (0.93)	Depression (1.47)	Self-harm (1.51)	Diabetes (1.33)	Stroke (0.58)

Ratio of observed to expected DALYs
(0.0-0.66)
(0.67-0.83)
(0.84-0.95)
(0.96-1.10)
(1.11-1.22)
(1.23-1.44)
(1.45-1.81)
(1.82-2.56)
>2.56

See Figure 7 caption for details. See Appendix Table 2 in Supplement 2 for explanation of terms.

Table 7. Percent Change in Age-Standardized Summary Exposure Values for the Leading 10 Risk Factors for the United States, the 50 States, and Washington, DC, 1990-2016, Both Sexes

	% Change (95% Uncertainty Interval)									
	High BMI	Smoking	High Fasting Plasma Glucose	High Systolic BP	Drug Use	Alcohol Use	High Total Cholesterol	Diet Low in Whole Grains	Impaired Kidney Function	Diet Low in Fruits
United States	53.2 (41.5 to 67.2)	−42.8 (−47.1 to −37.2)	76.0 (44.4 to 144.2)	−13.3 (−13.9 to −12.6)	10.1 (7.5 to 12.8)	6.0 (−24.2 to 42.1)	−17.2 (−19.4 to −15.4)	−8.0 (−12.5 to −0.8)	0.5 (−0.6 to 2.1)	−11.1 (−14.5 to −8.7)
Alabama	66.7 (48.0 to 94.3)	−24.3 (−33.0 to −15.2)	123.2 (53.3 to 289.8)	−12.7 (−15.8 to −9.7)	9.1 (5.6 to 11.8)	4.7 (−51.0 to 99.5)	−15.9 (−19.2 to −12.8)	−8.3 (−14.4 to −0.6)	2.9 (1.4 to 5.0)	4.4 (1.8 to 8.2)
Alaska	32.2 (19.0 to 49.3)	−32.0 (−41.2 to −22.9)	36.4 (7.8 to 101.6)	−12.0 (−14.8 to −9.1)	39.2 (26.5 to 53.5)	12.0 (−50.8 to 139.6)	−17.7 (−21.5 to −14.4)	−7.8 (−14.4 to −0.5)	2.6 (0.7 to 5.4)	−0.8 (−4.6 to 3.0)
Arizona	60.7 (40.6 to 88.9)	−48.4 (−54.5 to −40.8)	61.7 (22.9 to 164.9)	−12.6 (−15.4 to −9.6)	6.5 (3.2 to 9.3)	7.5 (−48.7 to 120.4)	−15.4 (−19.0 to −12.2)	−8.3 (−14.7 to −0.7)	1.0 (−0.5 to 3.3)	−33.5 (−41.2 to −28.0)
Arkansas	56.0 (36.8 to 80.3)	−27.3 (−34.6 to −19.4)	78.6 (30.1 to 205.2)	−12.7 (−15.7 to −9.7)	−0.9 (−4.3 to 1.8)	3.1 (−49.8 to 103.7)	−15.6 (−19.3 to −12.4)	−8.2 (−14.3 to −0.6)	3.9 (2.0 to 5.6)	−5.0 (−8.5 to −2.2)
California	54.4 (36.6 to 75.6)	−60.5 (−67.2 to −51.3)	42.6 (16.3 to 117.8)	−12.9 (−15.8 to −9.7)	11.3 (8.4 to 14.5)	10.6 (−43.1 to 124.6)	−16.1 (−19.6 to −13.1)	−7.9 (−14.2 to −0.6)	0.6 (−0.8 to 1.9)	−35.0 (−47.9 to −27.6)
Colorado	45.4 (29.9 to 64.8)	−48.2 (−54.9 to −40.6)	31.2 (10.7 to 78.0)	−13.5 (−16.3 to −10.6)	31.2 (19.1 to 47.5)	6.7 (−49.3 to 130.5)	−17.2 (−20.7 to −13.8)	−7.9 (−14.1 to −0.5)	−0.4 (−1.9 to 1.6)	−11.0 (−17.0 to −7.3)
Connecticut	53.6 (35.6 to 76.2)	−47.2 (−54.4 to −38.7)	83.6 (25.9 to 235.6)	−14.3 (−17.2 to −11.3)	4.4 (1.5 to 7.5)	4.6 (−49.6 to 109.9)	−18.3 (−21.9 to −15.1)	−7.4 (−13.5 to −0.5)	−0.5 (−2.7 to 2.1)	−9.5 (−16.3 to −5.2)
Delaware	40.0 (25.3 to 57.7)	−36.6 (−42.8 to −29.9)	96.5 (32.8 to 247.9)	−13.1 (−16.1 to −10.0)	7.3 (2.5 to 12.8)	8.8 (−48.3 to 119.5)	−17.1 (−20.8 to −13.8)	−8.0 (−14.3 to −0.6)	−0.8 (−3.1 to 1.9)	7.9 (4.5 to 12.9)
Florida	51.0 (34.9 to 71.4)	−47.6 (−54.4 to −39.0)	100.2 (37.9 to 249.5)	−12.3 (−15.3 to −9.2)	16.9 (12.6 to 21.6)	5.4 (−46.1 to 117.0)	−16.8 (−20.3 to −13.8)	−8.1 (−14.8 to −0.5)	3.0 (1.3 to 5.0)	−0.0 (−3.5 to 3.9)
Georgia	57.5 (38.9 to 81.5)	−43.1 (−50.4 to −34.1)	110.5 (47.0 to 273.2)	−13.7 (−16.7 to −10.2)	15.2 (11.7 to 19.8)	−8.9 (−57.4 to 75.8)	−17.4 (−20.9 to −14.1)	−8.3 (−14.8 to −0.7)	0.7 (−0.8 to 2.9)	−20.5 (−28.2 to −15.4)
Hawaii	47.5 (30.3 to 67.1)	−36.2 (−43.4 to −27.3)	64.8 (22.8 to 181.6)	−14.9 (−17.9 to −12.0)	29.7 (17.8 to 44.8)	9.7 (−47.5 to 114.3)	−15.1 (−18.6 to −11.9)	−8.0 (−14.2 to −0.5)	2.7 (1.1 to 4.5)	7.7 (4.4 to 13.1)
Idaho	61.9 (39.7 to 90.2)	−29.5 (−38.1 to −20.8)	71.5 (22.5 to 201.7)	−12.8 (−15.9 to −9.8)	3.7 (−2.1 to 10.8)	2.9 (−53.0 to 99.2)	−16.4 (−20.0 to −13.3)	−8.3 (−14.7 to −0.7)	3.5 (1.7 to 5.4)	4.8 (2.5 to 8.2)
Illinois	48.6 (33.2 to 69.5)	−47.2 (−53.1 to −39.3)	56.0 (18.8 to 166.8)	−13.6 (−16.6 to −10.6)	16.9 (12.1 to 22.1)	7.6 (−47.7 to 118.1)	−17.2 (−20.8 to −14.4)	−7.9 (−14.2 to −0.5)	−1.1 (−2.8 to 1.0)	7.1 (4.3 to 11.3)
Indiana	55.4 (37.6 to 77.3)	−25.4 (−34.3 to −15.7)	116.7 (45.3 to 265.0)	−12.6 (−15.5 to −9.5)	8.5 (4.8 to 11.8)	4.2 (−49.4 to 118.6)	−17.4 (−20.8 to −14.3)	−8.2 (−15.0 to −0.5)	2.9 (1.3 to 5.0)	−2.4 (−5.5 to 0.1)
Iowa	57.2 (37.4 to 82.6)	−26.6 (−34.2 to −18.3)	59.4 (20.2 to 151.8)	−13.2 (−16.0 to −10.2)	40.8 (32.6 to 48.5)	3.0 (−52.7 to 116.9)	−17.7 (−21.0 to −14.4)	−7.9 (−13.9 to −0.5)	2.4 (0.9 to 4.1)	−8.6 (−13.2 to −5.4)
Kansas	51.7 (34.1 to 73.9)	−29.4 (−38.2 to −19.2)	81.4 (27.2 to 235.4)	−12.7 (−15.6 to −9.6)	32.7 (25.0 to 40.6)	4.3 (−52.2 to 111.6)	−17.1 (−20.6 to −13.6)	−8.0 (−14.2 to −0.5)	3.2 (1.5 to 5.0)	−4.9 (−8.5 to −2.2)
Kentucky	57.0 (38.0 to 79.2)	−19.8 (−26.8 to −12.9)	115.8 (49.5 to 281.7)	−12.5 (−15.6 to −9.5)	11.2 (5.2 to 18.6)	7.9 (−45.9 to 116.6)	−16.9 (−20.7 to −13.9)	−8.2 (−14.6 to −0.7)	4.5 (2.4 to 6.4)	−4.3 (−7.5 to −1.7)
Louisiana	52.0 (36.8 to 71.6)	−31.9 (−38.7 to −24.7)	108.0 (40.0 to 277.3)	−12.7 (−15.5 to −9.5)	4.5 (1.2 to 7.3)	10.6 (−47.2 to 132.5)	−15.3 (−18.7 to −11.9)	−8.0 (−14.2 to −0.5)	2.6 (1.1 to 4.5)	−5.3 (−9.4 to −2.5)
Maine	47.9 (32.1 to 68.4)	−29.1 (−35.8 to −22.0)	81.2 (26.7 to 217.5)	−13.1 (−16.1 to −10.1)	9.5 (5.8 to 12.8)	10.3 (−49.5 to 134.0)	−16.9 (−20.4 to −13.5)	−8.2 (−14.5 to −0.6)	1.8 (0.3 to 3.9)	−19.7 (−26.4 to −15.2)
Maryland	48.1 (32.7 to 66.3)	−51.5 (−57.2 to −44.7)	98.9 (32.4 to 247.8)	−13.8 (−16.8 to −10.6)	−9.7 (−15.0 to −5.2)	6.8 (−50.4 to 125.4)	−18.2 (−21.8 to −15.1)	−7.7 (−13.6 to −0.4)	−3.6 (−6.5 to −0.4)	−57.3 (−69.6 to −48.7)
Massachusetts	52.0 (34.0 to 75.4)	−46.0 (−52.3 to −38.7)	45.3 (16.0 to 129.6)	−14.1 (−16.9 to −11.1)	2.5 (−1.0 to 6.1)	9.0 (−49.2 to 123.2)	−19.6 (−23.1 to −16.7)	−7.6 (−13.8 to −0.4)	−0.2 (−1.6 to 1.3)	−7.2 (−13.1 to −3.0)
Michigan	41.9 (28.3 to 60.0)	−37.3 (−43.9 to −30.0)	114.8 (46.6 to 276.7)	−13.4 (−16.3 to −10.1)	13.7 (10.0 to 16.8)	5.6 (−49.8 to 112.7)	−17.5 (−21.4 to −14.4)	−8.3 (−14.7 to −0.7)	0.9 (−0.5 to 2.9)	2.5 (−1.2 to 7.0)
Minnesota	52.9 (36.2 to 74.8)	−37.3 (−44.1 to −28.6)	48.0 (19.1 to 140.3)	−14.0 (−17.0 to −10.9)	23.5 (16.2 to 31.9)	5.6 (−48.3 to 115.5)	−18.3 (−21.5 to −15.2)	−7.8 (−13.9 to −0.4)	2.2 (0.7 to 4.1)	−4.1 (−8.2 to −0.9)
Mississippi	54.3 (37.1 to 77.0)	−20.3 (−30.3 to −10.3)	127.2 (56.6 to 298.8)	−11.8 (−15.1 to −8.6)	7.0 (4.4 to 10.0)	8.8 (−45.3 to 115.6)	−14.8 (−18.3 to −11.8)	−8.2 (−14.6 to −0.7)	4.1 (2.2 to 6.1)	0.1 (−2.9 to 2.9)
Missouri	54.3 (35.8 to 78.0)	−29.6 (−36.7 to −22.2)	111.8 (44.2 to 271.5)	−12.6 (−15.6 to −9.5)	42.8 (38.1 to 46.7)	5.4 (−50.3 to 122.2)	−16.4 (−19.7 to −13.1)	−8.1 (−14.2 to −0.7)	2.5 (1.0 to 4.6)	−2.2 (−5.5 to 0.7)

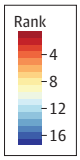
(continued)

Table 7. Percent Change in Age-Standardized Summary Exposure Values for the Leading 10 Risk Factors for the United States, the 50 States, and Washington, DC, 1990-2016, Both Sexes (continued)

	% Change (95% Uncertainty Interval)									
	High BMI	Smoking	High Fasting Plasma Glucose	High Systolic BP	Drug Use	Alcohol Use	High Total Cholesterol	Diet Low in Whole Grains	Impaired Kidney Function	Diet Low in Fruits
Montana	53.6 (36.9 to 77.2)	−26.5 (−35.5 to −16.0)	40.6 (14.9 to 116.1)	−12.9 (−15.5 to −10.1)	22.0 (15.1 to 29.2)	8.1 (−50.3 to 127.1)	−16.8 (−20.3 to −13.4)	−8.0 (−14.2 to −0.5)	0.0 (−1.6 to 2.3)	−5.0 (−8.9 to −2.1)
Nebraska	57.1 (39.5 to 80.5)	−30.1 (−37.2 to −22.1)	88.0 (31.4 to 229.0)	−13.6 (−16.4 to −10.6)	25.2 (17.5 to 33.5)	4.0 (−51.2 to 116.2)	−17.6 (−21.1 to −14.4)	−7.8 (−13.9 to −0.6)	2.3 (0.9 to 4.3)	−9.2 (−14.2 to −5.8)
Nevada	36.7 (21.5 to 54.2)	−51.2 (−57.7 to −43.7)	81.1 (23.6 to 228.6)	−12.2 (−15.2 to −9.4)	14.6 (11.7 to 17.9)	5.9 (−51.6 to 122.6)	−17.6 (−21.0 to −14.4)	−8.2 (−14.7 to −0.6)	0.7 (−1.0 to 3.1)	6.2 (2.7 to 10.9)
New Hampshire	57.7 (38.1 to 83.6)	−37.0 (−42.9 to −29.5)	79.5 (25.7 to 210.3)	−13.0 (−16.1 to −10.0)	2.2 (−2.0 to 7.1)	14.0 (−47.1 to 142.8)	−17.8 (−21.3 to −14.5)	−7.8 (−14.0 to −0.5)	0.4 (−1.4 to 3.0)	−3.5 (−7.8 to 0.1)
New Jersey	48.0 (32.4 to 67.8)	−52.5 (−59.1 to −44.7)	75.4 (24.4 to 198.9)	−14.6 (−17.5 to −11.6)	9.2 (5.5 to 13.5)	3.5 (−52.0 to 115.9)	−18.2 (−22.1 to −15.0)	−7.6 (−13.8 to −0.4)	−2.4 (−4.5 to 0.1)	−2.7 (−6.6 to 0.9)
New Mexico	68.2 (45.6 to 99.7)	−37.5 (−45.6 to −29.3)	74.3 (27.3 to 190.2)	−11.7 (−14.7 to −8.6)	23.8 (19.0 to 29.2)	10.4 (−46.9 to 121.0)	−15.3 (−18.8 to −12.0)	−8.2 (−14.4 to −0.7)	3.5 (1.6 to 5.6)	−10.2 (−15.7 to −6.5)
New York	51.3 (34.6 to 72.4)	−51.4 (−58.5 to −42.4)	105.4 (42.2 to 262.3)	−14.0 (−17.1 to −11.0)	1.3 (−1.3 to 3.9)	4.5 (−48.4 to 115.5)	−18.6 (−22.1 to −15.6)	−7.8 (−14.1 to −0.5)	−2.3 (−4.7 to 0.5)	−5.8 (−10.6 to −2.4)
North Carolina	54.1 (38.2 to 73.7)	−37.1 (−44.6 to −28.8)	109.2 (42.2 to 272.4)	−13.9 (−17.1 to −10.6)	14.2 (9.2 to 19.6)	6.9 (−46.7 to 105.5)	−17.4 (−20.9 to −14.3)	−8.2 (−14.3 to −0.6)	1.6 (0.1 to 3.6)	−6.5 (−10.5 to −3.6)
North Dakota	60.8 (41.8 to 87.7)	−19.4 (−29.4 to −8.6)	68.6 (23.9 to 190.9)	−13.8 (−16.6 to −10.9)	71.3 (58.2 to 86.3)	7.1 (−49.9 to 118.2)	−17.7 (−21.6 to −14.4)	−7.5 (−13.6 to −0.4)	1.9 (0.4 to 3.6)	−24.8 (−33.8 to −19.2)
Ohio	49.6 (34.0 to 70.6)	−28.1 (−35.0 to −20.5)	118.3 (50.1 to 297.7)	−12.5 (−15.5 to −9.3)	12.1 (8.3 to 15.6)	9.2 (−49.2 to 118.8)	−17.1 (−20.8 to −13.9)	−8.1 (−14.5 to −0.5)	2.2 (0.8 to 4.5)	−37.2 (−50.5 to −29.4)
Oklahoma	69.3 (47.7 to 99.6)	−26.6 (−33.5 to −19.3)	118.0 (50.9 to 292.2)	−13.2 (−15.9 to −10.0)	12.1 (8.2 to 16.6)	1.8 (−49.9 to 113.6)	−15.1 (−18.6 to −12.1)	−8.0 (−14.0 to -0.6)	3.7 (1.8 to 5.6)	−5.2 (−9.2 to −2.4)
Oregon	48.1 (33.0 to 67.0)	−44.0 (−50.5 to −36.0)	40.9 (14.1 to 110.0)	−13.7 (−16.8 to −10.5)	−0.2 (−3.8 to 3.0)	5.5 (−47.6 to 119.8)	−16.9 (−20.4 to −13.7)	−8.1 (−14.4 to −0.6)	1.9 (0.6 to 3.8)	−8.1 (−12.7 to −4.8)
Pennsylvania	46.4 (28.9 to 66.0)	−38.6 (−44.9 to −31.4)	1.7 (−36.4 to 50.1)	−13.5 (−16.6 to −10.4)	12.9 (9.4 to 16.5)	6.9 (−48.0 to 113.8)	−17.5 (−21.3 to −14.3)	−7.9 (−14.2 to −0.6)	−1.2 (−3.0 to 0.9)	−7.8 (−13.2 to −4.5)
Rhode Island	58.2 (40.4 to 78.5)	−37.7 (−46.2 to −28.8)	87.5 (32.7 to 230.9)	−13.2 (−16.4 to −10.2)	2.9 (0.7 to 5.0)	7.1 (−49.2 to 123.8)	−16.9 (−20.6 to −13.7)	−7.9 (−14.2 to −0.5)	−0.4 (−2.1 to 1.9)	0.6 (−2.8 to 4.5)
South Carolina	59.7 (41.9 to 83.9)	−32.1 (−40.6 to −22.7)	123.2 (54.8 to 304.8)	−13.1 (−16.2 to −10.0)	2.1 (−0.8 to 4.6)	9.3 (−46.5 to 126.1)	−16.4 (−19.8 to −13.4)	−8.2 (−14.4 to −0.6)	2.3 (0.8 to 4.6)	4.1 (1.5 to 7.4)
South Dakota	59.0 (38.7 to 88.1)	−21.5 (−31.4 to −10.1)	51.1 (21.6 to 129.6)	−13.2 (−16.1 to −10.1)	75.4 (60.2 to 92.0)	4.6 (−49.4 to 131.0)	−17.3 (−20.8 to −14.1)	−7.7 (−14.0 to −0.4)	1.9 (0.6 to 3.9)	−38.8 (−52.8 to −30.6)
Tennessee	64.3 (46.4 to 88.2)	−25.1 (−32.5 to −17.6)	110.4 (44.7 to 269.2)	−12.8 (−15.9 to −9.8)	6.6 (3.2 to 9.7)	5.9 (−47.4 to 105.7)	−15.9 (−19.3 to −12.8)	−8.0 (−14.6 to −0.5)	3.2 (1.5 to 5.1)	−8.1 (−12.2 to −5.2)
Texas	62.5 (42.9 to 86.9)	−51.2 (−57.7 to −42.5)	95.1 (34.5 to 251.3)	−13.4 (−16.5 to -10.3)	19.6 (12.9 to 27.1)	6.9 (−49.5 to 121.1)	−18.0 (−21.7 to −14.9)	−8.0 (−14.0 to -0.6)	3.8 (2.1 to 5.1)	−10.3 (−15.9 to −7.0)
Utah	59.4 (40.6 to 84.1)	−34.9 (−44.5 to −23.2)	72.4 (21.3 to 199.0)	−12.9 (−15.8 to −9.9)	1.4 (−1.9 to 4.2)	3.8 (−45.6 to 96.4)	−16.8 (−20.5 to −13.4)	−8.2 (−14.5 to −0.6)	2.1 (0.7 to 3.7)	−5.7 (−9.8 to −2.7)
Vermont	47.8 (32.9 to 69.8)	−29.8 (−37.5 to −21.3)	38.6 (15.6 to 103.5)	−12.9 (−16.0 to −10.0)	5.3 (−2.3 to 14.6)	14.3 (−46.9 to 146.4)	−17.2 (−20.6 to −14.0)	−7.8 (−14.3 to −0.5)	−2.4 (−5.0 to 0.6)	−2.3 (−6.1 to 0.7)
Virginia	54.3 (37.1 to 76.3)	−44.4 (−50.3 to −37.6)	96.8 (33.7 to 247.8)	−14.4 (−17.5 to −11.5)	13.9 (9.1 to 19.3)	9.0 (−46.6 to 123.0)	−17.9 (−21.4 to −14.8)	−7.8 (−13.9 to −0.5)	1.0 (−0.6 to 3.3)	−6.6 (−11.2 to −3.1)
Washington	56.9 (39.5 to 79.7)	−47.9 (−54.2 to −40.1)	44.8 (18.3 to 113.4)	−13.6 (−16.4 to −10.5)	16.0 (13.3 to 18.8)	5.2 (−48.7 to 116.7)	−18.0 (−21.8 to −15.0)	−7.9 (−14.0 to −0.5)	1.6 (0.3 to 3.6)	−7.7 (−12.6 to −4.4)
West Virginia	53.7 (38.4 to 75.4)	−11.2 (−18.6 to −2.3)	129.2 (51.7 to 301.8)	−13.1 (−16.2 to −10.1)	10.3 (6.5 to 13.7)	5.4 (−47.1 to 100.6)	−16.7 (−20.2 to −13.6)	−8.3 (−14.6 to −0.7)	5.2 (2.8 to 7.2)	−7.0 (−11.3 to −4.1)
Wisconsin	52.3 (36.7 to 72.8)	−32.6 (−39.8 to −24.4)	53.2 (15.7 to 156.5)	−13.1 (−16.0 to −9.8)	15.4 (10.3 to 21.6)	5.6 (−50.3 to 114.4)	−17.5 (−21.3 to −14.0)	−8.0 (−14.3 to −0.5)	2.4 (0.9 to 4.3)	−4.8 (−8.9 to −1.5)
Wyoming	52.3 (34.4 to 76.6)	−31.9 (−39.9 to −22.6)	41.1 (16.8 to 109.1)	−13.7 (−16.6 to −10.8)	116.0 (91.2 to 141.9)	10.3 (−49.4 to 146.5)	−17.7 (−21.3 to −14.5)	−7.5 (−13.8 to −0.4)	1.2 (−0.2 to 3.4)	−3.9 (−7.7 to −0.8)
Washington, DC	46.3 (31.2 to 65.0)	−51.5 (−59.6 to −40.3)	41.1 (7.3 to 120.5)	−15.6 (−18.6 to −12.5)	9.3 (5.8 to 15.9)	0.4 (−56.0 to 112.8)	−18.8 (−22.4 to −15.2)	−7.3 (−13.2 to −0.4)	−10.3 (−13.7 to −5.8)	5.6 (1.1 to 12.2)

Figure 10. Ranking of Risk Factors in 2016 for the United States Overall, the 50 States, and the District of Columbia According to the Number of Disability-Adjusted Life-Years Related to Each Risk Factor

	Tobacco use	High body mass index	Dietary risks	Alcohol and drug use	High fasting plasma glucose	High systolic blood pressure	High total cholesterol	Impaired kidney function	Occupational risks	Air pollution	Low physical activity	Child and maternal malnutrition	Low bone mineral density	Unsafe sex	Sexual abuse and violence	Residential radon and lead exposure	Unsafe water, sanitation, and handwashing
United States	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Alabama	1	3	2	6	4	5	7	8	9	10	11	12	13	14	16	15	17
Alaska	2	3	4	1	5	6	7	9	8	12	11	10	14	15	13	16	17
Arizona	3	2	4	1	5	6	7	9	8	10	11	12	13	15	14	16	17
Arkansas	1	3	2	6	4	5	7	8	9	10	11	12	13	14	16	15	17
California	4	1	3	2	5	6	7	8	9	10	11	12	13	14	15	16	17
Colorado	2	3	4	1	6	5	7	9	8	10	12	11	13	15	14	16	17
Connecticut	2	1	4	3	5	6	7	9	8	10	11	12	13	14	15	16	17
Delaware	1	2	3	5	4	6	7	8	9	10	11	12	14	13	15	16	17
District of Columbia	4	2	3	1	6	5	7	8	9	12	13	11	16	10	15	14	17
Florida	1	2	3	5	4	6	7	8	9	10	11	13	14	12	15	16	17
Georgia	1	2	3	4	5	6	7	8	9	10	11	12	14	13	15	16	17
Hawaii	3	1	2	5	4	6	7	8	9	12	10	11	13	14	15	16	17
Idaho	1	3	2	4	5	6	7	9	8	10	11	12	13	16	14	15	17
Illinois	2	1	3	4	6	5	7	9	8	10	11	12	13	14	16	15	17
Indiana	1	2	3	5	4	6	7	9	8	10	11	12	13	15	14	16	17
Iowa	1	3	2	6	4	5	7	9	8	10	11	13	12	16	15	14	17
Kansas	1	3	2	5	4	6	7	9	8	10	11	12	13	16	14	15	17
Kentucky	1	3	2	4	5	6	7	9	8	10	11	12	13	16	15	14	17
Louisiana	1	3	2	5	4	6	7	8	9	10	11	12	14	13	16	15	17
Maine	1	2	3	5	4	6	7	9	8	10	11	12	13	16	14	15	17
Maryland	2	1	3	5	4	6	7	8	9	10	12	11	14	13	15	16	17
Massachusetts	2	3	4	1	5	6	7	9	8	10	11	12	13	14	15	16	17
Michigan	1	2	3	5	4	6	7	8	9	10	11	12	13	16	15	14	17
Minnesota	1	2	3	4	5	6	8	9	7	10	11	12	13	16	14	15	17
Mississippi	1	3	2	6	4	5	7	8	9	10	11	12	14	13	16	15	17
Missouri	1	3	2	6	4	5	7	8	9	10	11	12	13	15	14	16	17
Montana	1	3	2	4	5	6	7	9	8	10	11	13	12	16	14	15	17
Nebraska	1	2	3	6	4	5	7	9	8	10	11	12	13	16	14	15	17
Nevada	1	3	4	2	5	6	7	9	8	10	11	12	14	15	13	16	17
New Hampshire	1	3	4	2	5	6	7	9	8	10	11	13	12	16	14	15	17
New Jersey	3	1	2	5	4	6	7	8	9	10	11	12	14	13	16	15	17
New Mexico	4	2	3	1	5	6	7	8	9	11	10	12	13	15	14	16	17
New York	3	1	2	5	4	6	7	8	9	10	11	13	14	12	15	16	17
North Carolina	1	3	2	5	4	6	7	8	9	10	11	12	13	14	15	16	17
North Dakota	2	1	3	5	4	6	7	9	8	12	10	11	13	16	14	15	17
Ohio	1	2	3	5	4	6	7	9	8	10	11	12	13	15	16	14	17
Oklahoma	1	3	2	5	4	6	7	8	9	10	11	12	13	15	14	16	17
Oregon	1	3	4	2	5	6	8	9	7	10	11	13	12	15	14	16	17
Pennsylvania	1	2	3	4	5	6	7	9	8	10	11	12	13	14	16	15	17
Rhode Island	1	2	4	3	5	6	7	8	9	10	11	12	13	14	16	15	17
South Carolina	1	3	2	5	4	6	7	8	9	10	11	12	14	13	15	16	17
South Dakota	1	2	3	6	4	5	7	9	8	10	11	12	13	16	15	14	17
Tennessee	1	3	2	4	5	6	7	9	8	10	11	12	13	14	15	16	17
Texas	3	1	2	5	4	6	7	9	8	10	11	12	14	13	16	15	17
Utah	6	2	3	1	4	5	8	9	7	11	12	10	14	16	13	15	17
Vermont	1	3	2	4	5	6	7	9	8	10	11	13	12	16	14	15	17
Virginia	2	1	3	5	4	6	7	8	9	10	11	12	13	14	15	16	17
Washington	3	2	4	1	5	6	7	9	8	10	11	12	13	15	14	16	17
West Virginia	1	3	2	5	4	6	7	9	8	10	11	12	13	15	16	14	17
Wisconsin	1	2	3	4	5	6	7	9	8	10	11	12	13	16	14	15	17
Wyoming	1	3	4	2	6	5	7	9	8	10	11	12	13	16	14	15	17



recently garnered attention. The increases in these causes of death have been underway for 20 years, but the US health policy community has been slow to recognize this rising set of problems.

This study shows that high BMI, smoking, and high fasting plasma glucose are the 3 most important risk factors in the United States, and that although smoking is decreasing, BMI

and fasting plasma glucose levels are steadily increasing. These 2 risk factors pose unique challenges in the United States given that unabated, they have the potential to change the health trajectory for individuals in many states. Levels of overweight and obesity increased during the study period. US residents need to do more to maintain their weight or reduce it, when needed, as well as access systems to support these intentions.[57] Although physical activity increased during the study period, the levels of increase were not enough to control weight gain.[8] Physical inactivity is a risk factor for many diseases, but increasing activity is not enough on its own to reduce weight or prevent weight gain.[58] Obesity is associated with increased diabetes, cardiovascular diseases, some neoplasms, and poor health-related quality of life. This study calls for renewed efforts to control weight gain at the community level.

Several studies have reported on the effect of taxes on sugary drinks or subsidies to encourage consumption of healthy foods, although only a small fraction of obesity can be linked to sugar-sweetened beverages.[59] A comprehensive plan is needed to address obesity because it adversely affects health and drives use of health care resources.[4,60-62] Rising BMI is driving up fasting plasma glucose levels and diabetes, and diabetes increased as a cause of burden in almost all states during the study period. Diabetes is a costly disease that consumes approximately 4.82% of the US health care budget.[4] A recent study estimated that the cost of diabetes care increased by 6.1% from 1996 to 2013.[4] Diabetes is associated with many conditions and disabilities.[63] This rise in burden and its cost is noteworthy given the projected increase in diabetes as obesity increases in the United States.

This study shows that there were reductions in the death rates from cardiovascular diseases for all age groups. This progress has been, in part, influenced by reductions in systolic blood pressure and cholesterol, but the role of increased access to effective treatment has also been considerable.[64] In fact, age-standardized CVD death rates have decreased in the United States by 32.8% over the last 20 years. The important role of treatment in reducing death rates highlights the ongoing importance of ensuring financial and physical access to care and the importance of quality of care. As declines in the rate of CVD may be slowing down, adverse trends due to the diseases of despair and adverse risk trends may mean that historical progress in improved life expectancy may not continue in the future.

The strategies for dealing with the remaining inequalities and new threats are 3-fold: (1) address some of the key modifiable risks, including diet; tobacco, alcohol, and drug use; insufficient physical activity; and obesity; (2) improve access to and, more importantly, quality of care in key areas, such as chronic kidney disease and ongoing care for substance use disorders; and (3) address the social determinants of health. We have previously shown an association between socioeconomic and race/ethnicity factors and a 60% county-level variation in life expectancy, behavioral and metabolic risk factors and a 74% variation, and health care factors and a 27% variation.[65] Combined, these factors are associated with 74% of overall variation. We also reported that most of the association between socioeconomic and race/ethnicity factors and life expectancy was mediated through behavioral and metabolic risk factors. Research has shown that some environmental factors have an effect on risk factors such as obesity and low physical activity. Low socioeconomic settings often have an imbalance of few grocery stores and numerous fast food options, and access to safe outdoor places spaces for recreation is limited.[66,67]

To date, the strategies for addressing the social determinants of health in the United States have been elusive. Lack of progress and rising social inequalities should not engender complacency. Therefore, addressing risk factors may prove to be an important opportunity to reduce disparities and deal with some of the challenges for improving US health. Opportunities to decrease the burden of disease through reducing tobacco, alcohol, drug use, blood pressure, and cholesterol; increasing physical activity; and improving diet emphasize that the United States should invest more in prevention that targets these risks. To increase the likelihood of prevention to succeed, it has to be a priority for all stakeholders—physicians, nurses, hospital systems, policy makers, health insurance companies, patients and their families, and advocacy groups.

This study showed a wide range of challenges encountered by different states and by some counties within states. Given the diversity of risks, communities, and workplaces there is no simple menu of effective programs for risk reduction. Indeed, local experimentation to determine what works in a given community is likely to be necessary. There is a need to change strategies of funding and evaluation of innovative interventions and policies, and independent evaluation of whether these efforts work or not should be documented. To succeed, these innovative programs should forge a connection between health care provision and progress for individuals and communities in health outcomes. The notion of accountability should be broadened beyond providing high-quality care to encompass achieving risk reduction in partnership with patients and communities.

The interpersonal violence burden in the United States has to be properly addressed, although declines have been observed: the age-standardized death rate decreased 32.43% from 1990 to 2016 in the United States. However, self-harm by firearm accounted for 6.39 deaths per 100 000 persons in the United States in 2016, and physical violence by firearm accounted for 3.98 deaths per 100 000 persons. There is evidence that gun control achieved through background checks reduces homicide and suicide.[68-72] Previous studies reported success in reducing the burden of gun deaths through policy changes in Brazil and other countries.[71] Indeed, enforcing gun control policies has proven effective in reducing mortality in a variety of contexts. There is a need for comprehensive studies of the epidemiology of gun violence in the United States to inform the ongoing gun control debate.

Age-standardized death rates due to alcohol increased by 17.50% from 1990 to 2016 in the United States, and alcohol use disorders accounted for 2.89 deaths per 100 000 persons in 2016. Previous studies have shown that alcohol consumption and binge drinking have increased in the United States, especially among females.[6,73] Alcohol is a major risk factor for burden in the United States[10] and is associated with adverse outcomes including sexually transmitted diseases, violence, and accidents.[16] Traffic injuries have received a considerable amount of attention, but the true burden of alcohol is much bigger and goes beyond driving

while intoxicated. Programs to educate US residents about the true harms of excessive drinking are needed.

Many of the risk factors that contribute to the disparities in burden are amenable to medical treatment within the context of supportive behavioral and lifestyle changes. For example, many cardiovascular risk factors, such as high blood pressure and high cholesterol levels, can now be treated more effectively with early detection and proper follow-up. Safe, effective, and affordable antihypertensive medications are now widely available, often as generic preparations, especially at discount pharmacies, and in many cases, without the need for health insurance for medications. The Affordable Care Act (ACA) allows for the expansion of insurance coverage, which ultimately increases access to care. Indeed, the ACA expansion of Medicaid coverage in participating states to all nonelderly adults with incomes below 133% of the federal poverty level provides an opportunity for early detection and follow-up of some of the main health risk factors. However, many US residents do not have health insurance, even after the ACA was introduced, and hence have little access to medical diagnosis and treatment. Therefore, expanding health coverage for certain conditions and medications should be considered and adopted to reduce burden.

This study showed that the United States overall and many individual states have made progress in reducing mortality but have had limited success in reducing disability. For instance, the burden of drug use disorders in total DALYs increased in the United States during the study period by 61.4% and accounted for about 3.81 million DALYs in 2016, depressive disorders increased 17.32% and accounted for 2.72 million DALYs, and anxiety disorders increased by 16.7% and accounted for approximately 1.76 million DALYs in 2016. These findings point to an urgent need to address mental health and drug use disorders in the United States. There is a need for improved access to quality mental health care and screening to improve outcomes, as well as programs to prevent mental disorders and promote mental health.

As the US population ages, the burden of musculoskeletal disorders will increase. More US residents have neck and back pain, and the incidence of falls is increasing. Musculoskeletal disorders are associated with a high medical cost.[4] Preventive measures to reduce the burden of these risk factors in all stages of life are urgently needed. Programs for avoiding harm and injuries at work among both younger and older ages are needed. The programs should include prevention of falls in the older population through examining the risk factors that lead to falls among adults. Screening tools and interventions to address this burden should be implemented.

The results of GBD 2016 have shown that occupational risks and air pollution were the 9th and 10th leading causes for DALYs. Although the findings show reductions in attributable burden from 1990 to 2016, occupational risks still account for 948.75 DALYS per 100 000 persons, and air pollution accounts for 584.97 DALYs per 100 000 persons—large numbers in the United States. Indeed, renewed efforts to reduce the burden of environmental and occupational risks are needed to ensure continued progress in reducing their effect on health in the United States.

Several studies have shown that poor diet is a major challenge in the United States, and little improvement has occurred over the past decades. US residents are not consuming a healthy diet; they tend to consume more calories than needed, and composition is not ideal.[49,62] Some recent studies have shown modest improvement in certain aspects of US diet, especially decreases in consumption of sugary drinks. The United States needs a comprehensive program to improve dietary intake at national and local levels. This program should offer financial incentives and disincentives for more vs less healthful food products by agriculture producers, food manufacturers, and retailers, as well as for choices by consumers. This effort should also implement comprehensive wellness programs in schools, workplaces, and government offices, and inform the public of the importance of a healthful diet.

Limitations

Given the scope of this analysis, this study has several important limitations. The overall limitations of the GBD methods, as noted in other publications, apply to the US analysis. First, the accuracy of the estimates depends on the availability of data by time period and state. Second, it is challenging to separate measurement error from variation in disease occurrence. GBD corrects for known bias from nonreference methods or case definitions, but often has to rely on sparse data at the state level to make those adjustments. Third, GBD applies garbage code redistribution for 13% of causes of deaths in the United States; this ranged from 8.4% in South Dakota to 21.3% in Alabama. Therefore, the causes of death may not match those in other publications but are more robust because they control for the between-states variation in the prevalence of garbage codes. Fourth, GBD methods adjust for hospital admissions using a large nonrepresentative source of medical claims data. The generalizability of claims data, the use of primary diagnosis only or all diagnostic fields, and the trends of claims data have been questioned.[13] Also, there may be considerable interstate variation in how diseases are treated between inpatient and outpatient settings. GBD methods adjust for such potential biases by using a covariate on claims and hospital admissions data to correct for systematic error. Fifth, GBD includes risk-outcome pairs that meet the World Cancer Research Fund criteria of causality. However, some risk-outcome pairs might not meet criteria that develop as evidence from new studies is published. Sixth, there is limited information on dietary intake at the state level. The Behavioral Risk Factors Surveillance System has 6 dietary questions attempting to capture fruit and vegetable consumption. Therefore, GBD 2016 used commercial sales data to adjust estimates of dietary intake. Seventh, some of the data used in the analyses have a lower quality and consistency across states and age groups. GBD 2016 reports 95% UIs to show the effect of this limitation on the estimates. Eighth, the study reports disparities between states but does not examine the within-state variations of burden, which could be substantial, especially in large states. Ninth, claims data were only available through 2012 at the time of these analyses. Additionally, this study does not report the burden of the social determinants of health; it focuses only on behavioral, environmental, and metabolic risks.

Conclusions

There are wide differences in the burden of disease at the state level. Specific diseases and risk factors, such as drug use disorders, high BMI, poor diet, high fasting plasma glucose level, and alcohol use disorders are increasing and warrant increased attention. These data can be used to inform national health priorities for research, clinical care, and policy.

ARTICLE INFORMATION

Accepted for Publication: February 13, 2018.

US Burden of Disease Collaborators: Ali H. Mokdad, PhD; Katherine Ballestros, PhD; Michelle Echko, BS; Scott Glenn, MSc; Helen E. Olsen, MA; Erin Mullany, BA; Alex Lee, BS; Abdur Rahman Khan, MD; Alireza Ahmadi, MD; Alize J. Ferrari, PhD; Amir Kasaeian, PhD; Andrea Werdecker, PhD; Austin Carter, BS; Ben Zipkin, BS; Benn Sartorius, PhD; Berrin Serdar, PhD; Bryan L. Sykes, PhD; Chris Troeger, MPH; Christina Fitzmaurice, MD; Colin D. Rehm, PhD; Damian Santomauro, PhD; Daniel Kim, DrPH; Danny Colombara, PhD; David C. Schwebel, PhD; Derrick Tsoi, BS; Dhaval Kolte, MD; Elaine Nsoesie, PhD; Emma Nichols, BA; Eyal Oren, PhD; Fiona J. Charlson, PhD; George C. Patton, MD; Gregory A. Roth, MD; H. Dean Hosgood, PhD; Harvey A. Whiteford, PhD; Hmwe Kyu, PhD; Holly E. Erskine, PhD; Hsiang Huang, MD; Ira Martopullo, MPH; Jasvinder A. Singh, MD; Jean B. Nachega, PhD; Juan R. Sanabria, MD; Kaja Abbas, PhD; Kanyin Ong, PhD; Karen Tabb, PhD; Kristopher J. Krohn, MPH; Leslie Cornaby, BS; Louisa Degenhardt, PhD; Mark Moses, MHS; Maryam Farvid, PhD; Max Griswold, MA; Michael Criqui, MD; Michelle Bell, PhD; Minh Nguyen, BS; Mitch Wallin, MD; Mojde Mirarefin, MPH; Mostafa Qorbani, PhD; Mustafa Younis, DrPH; Nancy Fullman, MPH; Patrick Liu, MPH; Paul Briant, BS; Philimon Gona, PhD; Rasmus Havmoller, PhD; Ricky Leung, PhD; Ruth Kimokoti, MD; Shahrzad Bazargan-Hejazi, PhD; Simon I. Hay, DSc; Simon Yadgir, BS; Stan Biryukov, BS; Stein Emil Vollset, DrPH; Tahiya Alam, MPH; Tahvi Frank, BS; Talha Farid, MD; Ted Miller, PhD; Theo Vos, PhD; Till Bärnighausen, MD; Tsegaye Telwelde Gebrehiwot, MPH; Yuichiro Yano, MD; Ziyad Al-Aly, MD; Alem Mehari, MD; Alexis Handal, PhD; Amit Kandel, MBBS; Ben Anderson, MD; Brian Biroscak, PhD; Dariush Mozaffarian, MD; E. Ray Dorsey, MD; Eric L. Ding, ScD; Eun-Kee Park, PhD; Gregory Wagner, MD; Guoqing Hu, PhD; Honglei Chen, PhD; Jacob E. Sunshine, MD; Jagdish Khubchandani, PhD; Janet Leasher, OD; Janni Leung, PhD; Joshua Salomon, PhD; Jurgen Unutzer, MD; Leah Cahill, PhD; Leslie Cooper, MD; Masako Horino, MPH; Michael Brauer, ScD; Nicholas Breitborde, PhD; Peter Hotez, PhD; Roman Topor-Madry, PhD; Samir Soneji, PhD; Saverio Stranges, PhD; Spencer James, MD; Stephen Amrock, MD; Sudha Jayaraman, MD; Tejas Patel, MD; Tomi Akinyemiju, PhD; Vegard Skirbekk, PhD; Yohannes Kinfu, PhD; Zulfiqar Bhutta, PhD; Jost B. Jonas, MD; Christopher J. L. Murray, DPhil.

Affiliations of US Burden of Disease Collaborators: Institute for Health Metrics and Evaluation, University of Washington, Seattle (Mokdad, Ballestros, Echko, Glenn, Olsen, Mullany, Lee, Ferrari, Carter, Zipkin, Troeger, Fitzmaurice, Santomauro, Colombara, Tsoi, Nsoesie, Nichols, Charlson, Roth, Whiteford, Kyu, Erskine, Martopullo, Ong, Krohn, Cornaby, Degenhardt, Moses, Griswold, Nguyen, Mirarefin, Fullman, Liu, Briant, Hay, Yadgir, Biryukov, Vollset, Alam, Frank, Vos, Brauer, James, Murray); University of Louisville, Louisville, Kentucky (Khan, Farid); Kermanshah University of Medical Sciences, Kermanshah, Iran (Ahmadi); Karolinska Institutet, Stockholm, Sweden (Ahmadi, Havmoller); School of Public Health, University of Queensland, Brisbane, Australia (Ferrari, Santomauro, Charlson, Whiteford, Erskine); Queensland Centre for Mental Health Research, Brisbane, Australia (Ferrari, Santomauro, Charlson, Whiteford, Erskine); Hematology-Oncology and Stem Cell Transplantation Research Center, and Hematologic Malignancies Research Center, Tehran University of Medical Sciences, Tehran, Iran (Kasaeian); Competence Center Mortality-Follow-Up of the German National Cohort, Federal Institute for Population Research, Wiesbaden, Hessen, Germany (Werdecker); Public Health Medicine, School of Nursing and Public Health, University of KwaZulu-Natal, Durban, South Africa (Sartorius); UKZN Gastrointestinal Cancer Research Centre, South African Medical Research Council, Durban, South Africa (Sartorius); University of Colorado, Aurora (Serdar); Departments of Criminology, Law & Society, Sociology, and Public Health, University of California, Irvine (Sykes); Division of Hematology, Department of Medicine, University of Washington, Seattle, and Fred Hutchinson Cancer Research Center, Seattle (Fitzmaurice); Montefiore Medical Center, Bronx, New York (Rehm); Department of Health Sciences, Northeastern University, Boston, Massachusetts (Kim); University of Alabama at Birmingham (Schwebel, Singh, Akinyemiju); Division of Cardiology, Brown University, Providence, Rhode Island (Kolte); Division of Epidemiology & Biostatistics, Graduate School of Public Health, San Diego State University, San Diego, California (Oren); Murdoch Childrens Research Institute, Department of Paediatrics, University of Melbourne, Melbourne, Victoria, Australia (Patton); Albert Einstein College of Medicine, Bronx, New York (Hosgood); Cambridge Health Alliance, Cambridge, Massachusetts (Huang); Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania (Nachega); Stellenbosch University, Cape Town, Western Cape, South Africa (Nachega); Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland (Nachega); Joan C. Edwards School of Medicine, Marshall University, Huntington, West Virginia (Sanabria); Case Western Reserve University, Cleveland, Ohio (Sanabria); Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, England (Abbas); School of Social Work, University of Illinois at Urbana-Champaign (Tabb); National Drug and Alcohol Research Centre, University of New South Wales, Sydney, Australia (Degenhardt); Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts (Farvid, Ding, Cahill); Harvard/MGH Center on Genomics, Vulnerable Populations, and Health Disparities, Mongan Institute for Health Policy, Massachusetts General Hospital, Boston (Farvid); University of California, San Diego, La Jolla, California (Criqui); Yale University, New Haven, Connecticut (Bell, Biroscak); VA Medical Center, Washington, DC (Wallin); Neurology Department, Georgetown University, Washington, DC (Wallin); Hunger Action Los Angeles, Los Angeles, California (Mirarefin); Non-communicable Diseases Research Center, Alborz University of Medical Sciences, Karaj, Alborz, Iran (Qorbani); Jackson State University, Jackson, Mississippi (Younis); University of Massachusetts, Boston (Gona); State University of New York, Albany, Rensselaer, New York (R. Leung); Simmons College, Boston, Massachusetts (Kimokoti); College of Medicine, Charles R. Drew University of Medicine and Science, Los Angeles, California (Bazargan-Hejazi); David Geffen School of Medicine, University of California at Los Angeles (Bazargan-Hejazi); Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, United Kingdom (Hay); Center for Disease Burden, Norwegian Institute of Public Health, and Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway (Vollset); Pacific Institute for Research & Evaluation, Calverton, Maryland (Miller); School of Public Health, Curtin University, Perth, Australia (Miller); Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts (Bärnighausen, Salomon); Africa Health Research Institute, Mtubatuba, South Africa (Bärnighausen); Institute of Public Health, Heidelberg University, Heidelberg, Germany (Bärnighausen); Jimma University, Jimma, Oromia, Ethiopia (Gebrehiwot); Department of Preventive Medicine, Northwestern University, Chicago, Illinois (Yano); Washington University in St Louis, St Louis, Missouri (Al-Aly); College of Medicine, Howard University, Washington, DC (Mehari); University of New Mexico, Albuquerque (Handal); University at Buffalo, Buffalo, New York (Kandel); University of Washington, Seattle (Anderson, Sunshine, J. Leung, Unutzer); University of South Florida, Tampa (Biroscak); Friedman School of Nutrition Science and Policy, Tufts University, Boston, Massachusetts (Mozaffarian); University of Rochester Medical Center, Rochester, New York (Dorsey); Department of Medical Humanities and Social Medicine, College of Medicine, Kosin University, Busan, South Korea (Park); Department of Environmental Health, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts (Wagner); Department of Epidemiology and Health Statistics, School of Public Health, Central South University, Changsha, Hunan, China (Hu); Michigan State University, East Lansing (Chen); Department of Nutrition and Health Science, Ball State University, Muncie, Indiana (Khubchandani); College of Optometry, Nova Southeastern University, Fort Lauderdale, Florida (Leasher); School of Public Health, University of Queensland, Brisbane, Australia (J. Leung); Dalhousie University, Halifax, Canada (Cahill); Mayo Clinic, Jacksonville, Florida (Cooper); Bureau of Child, Family & Community Wellness, Nevada Division of Public and Behavioral Health, Carson City (Horino); University of British Columbia, Vancouver, Canada (Brauer); The Ohio

State University, Columbus (Breitborde); College of Medicine, Baylor University, Houston, Texas (Hotez); Institute of Public Health, Faculty of Health Sciences, Jagiellonian University Medical College, Kraków, Poland (Topor-Madry); Faculty of Health Sciences, Wroclaw Medical University, Wroclaw, Poland (Topor-Madry); Dartmouth College, Hanover, New Hampshire (Soneji); Department of Epidemiology & Biostatistics, Schulich School of Medicine & Dentistry, Western University, London, Canada (Stranges); Department of Population Health, Luxembourg Institute of Health, Strassen, Luxembourg (Stranges); Oregon Health & Science University, Portland (Amrock); Department of Surgery, Virginia Commonwealth University, Richmond (Jayaraman); White Plains Hospital, White Plains, New York (Patel); Norwegian Institute of Public Health, Oslo, Norway (Skirbekk); Columbia University, New York, New York (Skirbekk); Centre for Research and Action in Public Health, University of Canberra, Canberra, Australia (Kinfu); Centre of Excellence in Women and Child Health, Aga Khan University, Karachi, Pakistan (Bhutta); Centre for Global Child Health, The Hospital for Sick Children, Toronto, Canada (Bhutta); Department of Ophthalmology, Medical Faculty Mannheim, Ruprecht-Karls-University Heidelberg, Heidelberg, Germany (Jonas).

Author Contributions: Dr Murray had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Murray, Mokdad, Troeger, Bell, Vos, Sykes, Dorsey, Cooper, Ahmadi, Kinfu, Younis, James.
Acquisition, analysis, or interpretation of data: Murray, Mokdad, Olsen, Mullany, Echko, Glenn, Ballesteros, Sartorius, Serdar, Troeger, Fitzmaurice, Rehm, Santomauro, Hosgood, Patton, Roth, Kyu, Martopullo, Nichols, Charlson, Degenhardt, Farvid, Criqui, Mirarefin, Huang, Gona, Havmoeller, Nachega, Hay, Biryukov, Vollset, Miller, Vos, Bärnighausen, Yano, Biroscak, Mozaffarian, Ding, Hu, Leasher, J. Leung, Salomon, Unützer, Horino, Bhutta, Al-Aly, Nsoesie, Schwebel, Kasaeian, Werdecker, Sykes, Colombara, Tsoi, Kolte, Sunshine, R. Leung, Erskine, Krohn, Whiteford, Oren, Kim, Singh, Ferrari, Khan, Carter, Sanabria, Abbas, Ong, Tabb, Moses, Cornaby, Fullman, Liu, Bazargan-Hejazi, Kimokoti, Breitborde, Hotez, Brauer, Dorsey, Cooper, Topor-Madry, Stranges, Amrock, Jayaraman, Patel, Akinyemiju, Ahmadi, Nguyen, Wallin, Qorbani, Yadgir, Frank, Farid, Alam, Kandel, Anderson, Park, Wagner, Chen, Khubchandani, Cahill, Soneji, James, Gebrehiwot.
Drafting of the manuscript: Murray, Mokdad, Olsen, Mullany, Glenn, Ballesteros, Roth, Vos, Horino, Krohn, Sanabria, Fullman, Ahmadi, Frank.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Mokdad, Glenn, Troeger, Fitzmaurice, Rehm, Santomauro, Roth, Kyu, Martopullo, Nichols, Charlson, Hay, Biryukov, Vos, J. Leung, Kasaeian, Sykes, Colombara, Tsoi, Erskine, Ferrari, Carter, Ong, Moses, Cornaby, Liu, Nguyen, Qorbani, Yadgir, Frank, Khubchandani.
Obtained funding: Murray, Mokdad.
Administrative, technical, or material support: Murray, Mokdad, Olsen, Echko, Nachega, Hay, Mozaffarian, Unützer, R. Leung, Krohn, Whiteford, Singh, Akinyemiju, Ahmadi, Wallin, Alam, Park, James.
Supervision: Murray, Mokdad, Olsen, Mullany, Criqui, Bell, Vos, Salomon, Stranges, Ahmadi, Alam.
Other (created custom analyses, tables, and figures, and fact checked the analysis): Echko, Ballesteros, Lee, Glenn.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Singh reports receipt of grants and personal fees from Savient, Takeda, and Crealta/Horizon; personal fees from Regeneron, Merz, Iroko, Bioiberica, Allergan, UBM LLC, WebMD, and the American College of Rheumatology during the conduct of the study; personal fees from DINORA; and nonfinancial support from Outcomes Measures in Rheumatology outside the submitted work. Dr Degenhardt reports receipt of grants from Mundipharma, Indivior, and Seqirus outside the submitted work. Dr Bell reports receipt of grants from the US Environmental Protection Agency and the National Institutes of Health (NIH) during the conduct of the study, and honorarium and/or travel reimbursement from NIH, the *American Journal of Public Health*, Columbia University, Washington University, Statistical Methods and Analysis of Environmental Health Data Workshop, North Carolina State University, and Global Research Laboratory and Seoul National University. Dr Mozaffarian reports receipt of personal fees from Acasti Pharma, GOED, DSM, Haas Avocado Board, Nutrition Impact, Pollock Communications, Boston Heart Diagnostics, and Bunge; other for serving on a scientific advisory board from Omada Health and Elysium Health; a grant from NIH and the National Heart, Lung, and Blood Institute; a grant from the Bill and Melinda Gates Foundation; and royalties from UpToDate outside the submitted work.

Funding/Support: This study was supported in part by the Intramural Program of the National Institutes of Health, the National Institute of Environmental Health Sciences, and in part by the Bill and Melinda Gates Foundation.

Role of the Funder/Sponsor: The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contribution: We thank Ashkan Afshin for assistance with revisions to the risk factor section of this article.

REFERENCES

1. Dwyer-Lindgren L, Mackenbach JP, van Lenthe FJ, Flaxman AD, Mokdad AH. Diagnosed and undiagnosed diabetes prevalence by county in the US, 1999-2012. *Diabetes Care*. 2016;39(9):1556-1562.

2. Wang H, Schumacher AE, Levitz CE, Mokdad AH, Murray CJ. Left behind: widening disparities for males and females in US county life expectancy, 1985-2010. *Popul Health Metr*. 2013;11(1):8.

3. Case A, Deaton A. Mortality and morbidity in the 21st century. https://www.brookings.edu/bpea-articles/mortality-and-morbidity-in-the-21st-century/. Accessed October 19, 2017.

4. Dieleman JL, Baral R, Birger M, et al. US spending on personal health care and public health, 1996-2013. *JAMA*. 2016;316(24):2627-2646.

5. Wang H, Abajobir AA, Abate KH, et al; GBD 2016 Mortality Collaborators. Global, regional, and national under-5 mortality, adult mortality, age-specific mortality, and life expectancy, 1970-2016. *Lancet*. 2017;390(10100):1084-1150.

6. Dwyer-Lindgren L, Flaxman AD, Ng M, et al. Drinking patterns in US counties from 2002 to 2012. *Am J Public Health*. 2015;105(6):1120-1127.

7. Dwyer-Lindgren L, Mokdad AH, Srebotnjak T, et al. Cigarette smoking prevalence in US counties: 1996-2012. *Popul Health Metr*. 2014;12(1):5.

8. Dwyer-Lindgren L, Freedman G, Engell RE, et al. Prevalence of physical activity and obesity in US counties, 2001-2011. *Popul Health Metr*. 2013;11:7.

9. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. *JAMA*. 2004;291(10):1238-1245.

10. Murray CJL, Atkinson C, Bhalla K, et al; U.S. Burden of Disease Collaborators. The state of US health, 1990-2010. *JAMA*. 2013;310(6):591-608.

11. Arias E. *Changes in Life Expectancy by Race and Hispanic Origin in the United Staes, 2013-2014*. Atlanta, GA: National Center for Health Statistics; 2016.

12. National Center for Health Statistics. Healthy People 2020. Published November 6, 2015. https://www.healthypeople.gov/. Accessed October 19, 2017.

13. Naghavi M, Abajobir AA, Abbafati C, et al; GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016. *Lancet*. 2017;390(10100):1151-1210.

14. Vos T, Abajobir AA, Abate KH, et al; GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016. *Lancet*. 2017;390(10100):1211-1259.

15. Hay SI, Abajobir AA, Abate KH, et al; GBD 2016 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2016. *Lancet*. 2017;390(10100):1260-1344.

16. Gakidou E, Afshin A, Abajobir AA, et al; GBD 2016 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2016. *Lancet*. 2017;390(10100):1345-1422.

17. Fullman N, Barber RM, Abajobir AA, et al; GBD 2016 SDG Collaborators. Measuring progress and projecting attainment on the basis of past trends of the health-related sustainable development goals in 188 countries. *Lancet*. 2017;390(10100):1423-1459.

18. National Center for Health Statistics. National Health and Nutrition Examination Survey. Published October 17, 2017. https://www.cdc.gov/nchs/nhanes/index.htm. Accessed November 11, 2017.

19. Healthcare Cost and Utilization Project. SID database documentation. https://www.hcup-us.ahrq.gov/db/state/siddbdocumentation.jsp. Accessed November 11, 2017.

20. Lindrooth RC, Perraillon MC, Hardy RY, Tung GJ. Understanding the relationship between Medicaid expansions and hospital closures. *Health Aff (Millwood)*. 2018;37(1):111-120.

21. Stevens GA, Alkema L, Black RE, et al. Guidelines for accurate and transparent health estimates reporting: the GATHER statement. *Lancet*. 2016;388(10062):e19-e23.

22. National Center for National Health Statistics. Ambulatory health care data. Published October 10, 2017. https://www.cdc.gov/nchs/ahcd/index.htm. Accessed November 11, 2017.

23. National Center for Health Care Statistics. National hospital ambulatory medical care survey series. https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/42. Accessed November 11, 2017.

24. Agency for Healthcare Research and Quality. Medical expenditure panel survey 1996-2013. https://meps.ahrq.gov/mepsweb/. Accessed November 11, 2017.

25. Harvard Medical School. National Comorbidity Survey. https://www.hcp.med.harvard.edu/ncs/index.php. Accessed November 11, 2017.

26. National Institute on Alcohol Abuse and Alcoholism. National Epidemiologic Survey on Alcohol and Related Conditions. http://pubs.niaaa.nih.gov/publications/arh29-2/toc29-2.htm. November 11, 2017.

27. Substance Abuse and Mental Health Services Administration. Results from the 2015 National Survey on Drug Use and Heath. https://www.samhsa.gov/data/sites/default/files/NSDUH-DetTabs-2015/NSDUH-DetTabs-2015/NSDUH-DetTabs-2015.pdf. November 11, 2017.

28. United States Department of Agriculture. Nationwide food surveys, 1935-1998. https://www.ars.usda.gov. Accessed October 19, 2017.

29. Truven Health Analytics. MarketScan research data. https://marketscan.truvenhealth.com/marketscanportal/. Accessed October 19, 2017.

30. National Center for National Health Statistics. National Health Interview Survey. https://www.cdc.gov/nchs/nhis/index.htm. Accessed October 19, 2017.

31. National Center for Chronic Disease Prevention. Behavioral Risk Factor Surveillance System. https://www.cdc.gov/brfss/index.html. Accessed October 19, 2017.

32. Centers for Disease Control and Prevention. National Nortifiable Disease Surveillance System (NNDSS): August 2015. https://wwwn.cdc.gov/Nndss/. Accessed November 11, 2017.

33. National Center for National Health Statistics. National Vital Statistics System: multiple cause of death data file, 1980-2014. Published August 21, 2017. https://www.cdc.gov/nchs/nvss/deaths.htm. Accessed October 19, 2017.

34. US Census Bureau. Intercensal county estimates by age, sex, race: 1980-1989. https://www.census.gov/data/tables/time-series/demo/popest/1980s-county.html. Accessed October 19, 2017.

35. University of California Berkeley. Human Mortality Database: published July 2013. http://www.mortality.org/ Accessed October 19, 2017.

36. Naghavi M, Makela S, Foreman K, et al. Algorithms for enhancing public health utility of national causes-of-death data. *Popul Health Metr*. 2010;8:9.

37. Wang H, Naghavi M, Allen C, et al; GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015. *Lancet*. 2016;388(10053):1459-1544.

38. Flaxman AD, Vos T, Murray CJL, eds. *An Integrative Metaregression Framework for Descriptive Epidemiology*. Seattle, WA: University of Washington Press; 2015.

39. Murthy VH. Facing addiction in the United States. *JAMA*. 2017;317(2):133-134.

40. International ICF. National Evaluation of the Drug-Free Communities Support Program. https://obamawhitehouse.archives.gov/sites/default/files/DFC2014Interim%20ReportExecutiveSummaryFinal.pdf. Accessed March 23, 2018.

41. Global Health Data Exchange. GBD results tool. http://ghdx.healthdata.org/gbd-results-tool. Accessed March 9, 2018.

42. Rehm CD, Peñalvo JL, Afshin A, Mozaffarian D. Dietary intake among US adults, 1999-2012. *JAMA*. 2016;315(23):2542-2553.

43. Sharma B, Bruner A, Barnett G, Fishman M. Opioid use disorders. *Child Adolesc Psychiatr Clin N Am*. 2016;25(3):473-487.

44. Lipari RN, Hughes A. Trends in heroin use in the United States: 2002 to 2013. In: Substance Abuse and Mental Health Services Administration, ed. *The CBHSQ Report: April 23, 2015*. Rockville, MD. Substance Abuse and Mental Health Services Administration; 2013, http://www.ncbi.nlm.nih.gov/books/NBK343534/. Accessed October 19, 2017.

45. Miller NS, Sheppard LM. The role of the physician in addiction prevention and treatment. *Psychiatr Clin North Am*. 1999;22(2):489-505.

46. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain—United States, 2016. *MMWR Recomm Rep*. 2016;65(1):1-49.

47. Larochelle MR, Liebschutz JM, Zhang F, Ross-Degnan D, Wharam JF. Opioid prescribing after nonfatal overdose and association with repeated overdose: a cohort study. *Ann Intern Med*. 2016;164(1):1-9.

48. Canetto SS, Sakinofsky I. The gender paradox in suicide. *Suicide Life Threat Behav*. 1998;28(1):1-23.

49. Miller M, Hemenway D. Guns and suicide in the United States. *N Engl J Med*. 2008;359(10):989-991.

50. Mokdad AH, Bowman BA, Ford ES. The continuing epidemics of obesity and diabetes in the United States. *JAMA*. 2001;286(10):1195-1200.

51. Haskell WL, Lee I-M, Pate RR, et al; American College of Sports Medicine; American Heart Association. Physical activity and public health. *Circulation*. 2007;116(9):1081-1093.

52. Veerman JL, Sacks G, Antonopoulos N, Martin J. The impact of a tax on sugar-sweetened beverages on health and health care costs: a modelling study. *PLoS One*. 2016;11(4):e0151460.

53. Brown DW, Balluz LS, Giles WH, et al; Behavioral Risk Factor Surveillance System (BRFSS). Diabetes mellitus and health-related quality of life among older adults. *Diabetes Res Clin Pract*. 2004;65(2):105-115.

54. Li C, Ford ES, Zhao G. Prevalence and correlates of undiagnosed depression among US adults with diabetes. *Diabetes Res Clin Pract*. 2009;83(2):268-279.

55. Urban LE, Roberts SB, Fierstein JL, Gary CE, Lichtenstein AH. Sodium, saturated fat, and trans fat content per 1000 kilocalories. *Prev Chronic Dis*. 2014;11:E228.

56. Mokdad AH, Ford ES, Bowman BA, et al. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *JAMA*. 2003;289(1):76-79.

57. Dalen JE, Alpert JS, Goldberg RJ, Weinstein RS. The epidemic of the 20(th) century: coronary heart disease. *Am J Med*. 2014;127(9):807-812.

58. Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, et al. Inequalities in life expectancy among US counties, 1980 to 2014. *JAMA Intern Med*. 2017;177(7):1003-1011.

59. Siegel M, Ross CS, King C III. The relationship between gun ownership and firearm homicide rates in the United States, 1981-2010. *Am J Public Health*. 2013;103(11):2098-2105.

60. Vigdor ER, Mercy JA. Do laws restricting access to firearms by domestic violence offenders prevent intimate partner homicide? *Eval Rev*. 2006;30(3):313-346.

61. Campbell JC, Webster D, Koziol-McLain J, et al. Risk factors for femicide in abusive relationships. *Am J Public Health*. 2003;93(7):1089-1097.

62. Santaella-Tenorio J, Cerdá M, Villaveces A, Galea S. What do we know about the association between firearm legislation and firearm-related injuries? *Epidemiol Rev*. 2016;38(1):140-157.

63. Anestis MD, Anestis JC. Suicide rates and state laws regulating access and exposure to handguns. *Am J Public Health*. 2015;105(10):2049-2058.

64. Liu S, Siegel PZ, Brewer RD. Prevalence of alcohol-impaired driving. *JAMA*. 1997;277(2):122-125.

65. Ikeda N, Sapienza D, Guerrero R, et al. Control of hypertension with medication. *Bull World Health Organ*. 2014;92(1):10-19C.

66. Mozaffarian D, Afshin A, Benowitz NL, et al. AHA scientific statement population approaches to improve diet, physical activity, and smoking habits. Circulation. 2012;126(12):1514-1563.

67. Micha R, Peñalvo JL, Cudhea F, et al. Association between dietary factors and mortality from heart disease, stroke, and type 2 diabetes in the United States. *JAMA*. 2017;317(9):912-924.

68. Olives C, Myerson R, Mokdad AH, Murray CJL, Lim SS. Prevalence, awareness, treatment, and control of hypertension in United States counties, 2001-2009. *PLoS One*. 2013;8(4):e60308.

69. Roth GA, Nguyen G, Forouzanfar MH, Mokdad AH, Naghavi M, Murray CJL. Estimates of global and regional premature cardiovascular mortality in 2025. *Circulation*. 2015;132(13):1270-1282.

70. Rangel C. HR 3590, 111th Congress (2009-2010): Patient Protection and Affordable Care Act. https://www.congress.gov/bill/111th-congress/house-bill/3590. Accessed December 16, 2017.

71. Hill BJ. What is the meaning of health? *Am J Law Med*. 2012;38(2-3):445-470.

72. Center for Behavioral Health Statistics and Quality. Behavioral health trends in the United States. https://www.samhsa.gov/data/sites/default/files/NSDUH-MHDetTabs2014/NSDUH-MHDetTabs2014.htm. Accessed October 19, 2017.

73. Mokdad AH. The Behavioral Risk Factors Surveillance System: past, present, and future. *Annu Rev Public Health*. 2009;30:43-54.

A Spotlight on World Obesity Rates

This week in the United States, people will celebrate the Thanksgiving Day holiday with family traditions that usually include a very large meal. In anticipation of waist-band stretching feasts, we decided to see how the United States fares against other populations in terms of obesity. The U.S. comes in near the top, having the 6th highest obesity rate. *The World Factbook* (/library/publications/resources/the-world-factbook/index.html) data also show that there are clear regional trends, even though obesity rates vary around the world.

Of the 10 countries or territories with the highest obesity rates, 5 are in Oceania. Another 4 are in the Middle East. Rates on this chart reflect the percent of a country's adult population that has Body Mass Index (BMI) greater than or equal to 30.0, which is considered obese.

Country	Obesity Rate
1. American Samoa (US territory) (https://www.cia.gov/library/publications/the-world-factbook/geos/aq.html)	74.60
2. Tokelau (https://www.cia.gov/library/publications/the-world-factbook/geos/tl.html)	63.40
3. Tonga (https://www.cia.gov/library/publications/the-world-factbook/geos/tn.html)	56.00
4. Kiribati (https://www.cia.gov/library/publications/the-world-factbook/geos/kr.html)	50.60
5. Saudi Arabia (https://www.cia.gov/library/publications/the-world-factbook/geos/sa.html)	35.60
6. United States (https://www.cia.gov/library/publications/the-world-factbook/geos/us.html)	33.90
7. United Arab Emirates (https://www.cia.gov/library/publications/the-world-factbook/geos/ae.html)	33.70
8. Egypt (https://www.cia.gov/library/publications/the-world-factbook/geos/eg.html)	30.30

9. Kuwait (https://www.cia.gov/library/publications/the-world-factbook/geos/ku.html)	28.80
10. New Zealand (https://www.cia.gov/library/publications/the-world-factbook/geos/nz.html)	26.50

Of the 10 countries with the lowest obesity rates, 8 are in Asia, and the remaining 2 are in Africa.

Country	**Obesity Rate**
1. Vietnam (https://www.cia.gov/library/publications/the-world-factbook/geos/vm.html)	0.50
2. Laos (https://www.cia.gov/library/publications/the-world-factbook/geos/la.html)	1.20
3. Madagascar (https://www.cia.gov/library/publications/the-world-factbook/geos/ma.html)	2.10
4. Indonesia (https://www.cia.gov/library/publications/the-world-factbook/geos/id.html)	2.40
5. China (https://www.cia.gov/library/publications/the-world-factbook/geos/ch.html)	2.90
6. Japan (https://www.cia.gov/library/publications/the-world-factbook/geos/ja.html)	3.10
7. Korea, South (https://www.cia.gov/library/publications/the-world-factbook/geos/ks.html)	3.20
8. Eritrea (https://www.cia.gov/library/publications/the-world-factbook/geos/er.html)	3.30
9. Philippines (https://www.cia.gov/library/publications/the-world-factbook/geos/rp.html)	4.30
10. Singapore (https://www.cia.gov/library/publications/the-world-factbook/geos/sn.html)	6.90

Explore *The World Factbook* “People (https://www.cia.gov/library/publications/the-world-factbook/rankorder/rankorderguide.html)” section to see more country comparisons of social and health data.

Historical Document
Posted: Nov 19, 2012 12:46 PM
Last Updated: Apr 30, 2013 01:03 PM

FEATURED STORY ARCHIVE

- **2020** (/news-information/featured-story-archive/2020-featured-story-archive/index)
- **2019** (/news-information/featured-story-archive/2019-featured-story-archive/index)
- **2018** (/news-information/featured-story-archive/2018-featured-story-archive/index)
- **2017 (/news-information/featured-story-archive/2017-featured-story-archive/index)**
- **2016 (/news-information/featured-story-archive/2016-featured-story-archive/index)**
- **2015** (/news-information/featured-story-archive/2015-featured-story-archive/index)
- **2014** (/news-information/featured-story-archive/2014-featured-story-archive/index)
- **2013** (/news-information/featured-story-archive/2013-featured-story-archive/index)
- **2012** (/news-information/featured-story-archive/2012-featured-story-archive/index)
- **2011** (/news-information/featured-story-archive/2011-featured-story-archive/index)
- **2010** (/news-information/featured-story-archive/2010-featured-story-archive/index)
- **2009 (/news-information/featured-story-archive/2009-featured-story-archive/index)**
- **2008** (/news-information/featured-story-archive/2008-featured-story-archive/index.html)
- **2007** (/news-information/featured-story-archive/2007-featured-story-archive)

April 02, 2020 1 min read

FDA clears IND application for natural killer cell-based COVID-19 therapy

The FDA cleared an investigational new drug application for CYNK-001 for the treatment of adults with COVID-19 infection, according to the agent's manufacturer.

CYNK-001 (Celularity) is an investigational, cryopreserved, allogeneic natural killer cell therapy derived from placental hematopoietic stem cells.

"This investigational new drug application represents a significant step toward a potential treatment of patients infected with COVID-19 virus, which is spreading globally at unanticipated rates," **Robert Hariri, MD, PhD,** founder, chairman and CEO of Celularity, said in a company-issued statement.

"With our initial clinical study, we will gain an understanding of the impact CYNK-001 can have on patients recently diagnosed with COVID-19," Hariri added. "We are hopeful to contribute to flattening the COVID-19 curve, expanding on the promising early results we've seen in our clinical studies in devastating cancers to patients with coronavirus."

This FDA clearance will allow Celularity to begin a phase 1/phase 2 clinical trial of its natural killer cell therapy in up to 86 patients with COVID-19. The investigational treatment is believed to be the first immunotherapy to receive this clearance by the FDA for treatment of COVID-19, according to the statement.

"Studies have established that there is robust activation of [natural killer] cells during viral infection regardless of the virus class," **Xiaokui Zhang, PhD,** chief scientific officer of Celularity, said in the statement.

Zhang said CYNK-001 has demonstrated "a range of biological activities" that would allow its cellular therapy to potentially recognize and kill infected cells.

"These functions suggest that CYNK-001 could provide a benefit to patients with COVID-19 in terms of limiting SARS-CoV-2 replication and disease progression by eliminating the infected cells," he added.

The FDA approved an IND for CYNK-001 in January for the treatment of glioblastoma multiforme, a type of brain cancer, as previously reported by Healio.

fda cell therapy covid-19 covid infection natural killer cell

Med Hypotheses. 2020 Apr 22;140:109777. doi: 10.1016/j.mehy.2020.109777. Online ahead of print.

Innate immunity in COVID-19 patients mediated by NKG2A receptors, and potential treatment using Monalizumab, Cholroquine, and antiviral agents

Ahmed Yaqinuddin [1], Junaid Kashir [2]

Affiliations

PMID: 32344314 PMCID: PMC7194824 DOI: 10.1016/j.mehy.2020.109777

Free PMC article

Abstract

Following the outbreak of a novel coronavirus (SARS-CoV-2), studies suggest that the resultant disease (COVID-19) is more severe in individuals with a weakened immune system. Cytotoxic T-cells (CTLs) and Natural Killer (NK) cells are required to generate an effective immune response against viruses, functional exhaustion of which enables disease progression. Patients with severe COVID-19 present significantly lower lymphocyte, and higher neutrophil, counts in blood. Specifically, $CD8^+$ lymphocytes and NK cells were significantly reduced in cases of severe infection compared to patients with mild infection and healthy individuals. The NK group 2 member A (NKG2A) receptor transduces inhibitory signalling, suppressing NK cytokine secretion and cytotoxicity. Overexpression of NKG2A has been observed on $CD8^+$ and NK cells of COVID-19 infected patients compared to healthy controls, while NKG2A overexpression also functionally exhausts $CD8^+$ cells and NK cells, resulting in a severely compromised innate immune response. Blocking NKG2A on $CD8^+$ cells and NK cells in cancers modulated tumor growth, restoring $CD8^+$ T and NK cell function. A recently proposed mechanism via which SARS-CoV-2 overrides innate immune response of the host is by over-expressing NKG2A on CD^+ T and NK cells, culminating in functional exhaustion of the immune response against the viral pathogen. Monalizumab is an inhibiting antibody against NKG2A which can restore the function of CD8 + T and NK cells in cancers, successfully ceasing tumor progression with no significant side effects in Phase 2 clinical trials. We hypothesize that patients with severe COVID-19 have a severely compromised innate immune response and could be treated via the use of Monalizumab, interferon α, chloroquine, and other antiviral agents.

Keywords: COVID-19; Innate immunity; Monalizumab; NKG2A; SARS.

Related information

MedGen

LinkOut - more resources

Full Text Sources

ClinicalKey
Elsevier Science
Europe PubMed Central
PubMed Central

Miscellaneous

NCI CPTAC Assay Portal

Randomized Controlled Trial Med Sci Sports Exerc. 2005 Jan;37(1):57-62.
doi: 10.1249/01.mss.0000149808.38194.21.

Immune response to a 30-minute walk

David C Nieman [1], Dru A Henson, Melanie D Austin, Victor A Brown

Affiliations
PMID: 15632669 DOI: 10.1249/01.mss.0000149808.38194.21

Abstract

Purpose: To measure several components of immune changes related to walking 30 min with or without an exercise assist device compared with sitting.

Methods: Fifteen healthy and nonobese female subjects (37.5 +/- 3.1 yr of age) accustomed to regular walking were recruited and tested for aerobic power (VO(2max) 34.4 +/- 1.4 mL.kg(-1).min(-1)). During three randomly assigned 30-min test sessions, subjects functioned as their own controls and either sat in the laboratory, walked at approximately 60% VO(2max), or walked at the same treadmill speed using the BODY BAT Aerobic Exerciser. This exercise assist device resembles a pair of baseball bats seamlessly joined together and is held with both hands and swung to shoulder height across the body in a natural side to side pendulum motion. Saliva and blood samples were collected pre- and postexercise, and 1 h postexercise, with the data statistically analyzed using a 3 x 3 repeated measures ANOVA.

Results: Walking with the exercise assist device increased oxygen consumption 11 +/- 2% and heart rate 8 +/- 2 beats.min(-1). The pattern of increase in blood counts for neutrophils, lymphocytes, monocytes, and natural killer cells, plasma interleukin-6 concentration, and PHA-induced lymphocyte proliferation differed significantly when comparing walking with sitting, but no differences were found between walking with or without the exercise assist device. No significant increases over time or interaction effects were measured for plasma cortisol concentration, salivary IgA output, or plasma interleukin-1 receptor antagonist concentration.

Conclusions: The use of an exercise assist device increased oxygen consumption 11% during walking, but did not alter the pattern of change in several components of immunity measured during walking alone in comparison to sitting. Walking caused modest and short-lived changes in immune parameters, most notably for neutrophil and natural killer blood cell counts.

Related information

MedGen
PubChem Compound
PubChem Compound (MeSH Keyword)
PubChem Substance

LinkOut - more resources

Full Text Sources

Ovid Technologies, Inc.
Wolters Kluwer

Research Materials

NCI CPTC Antibody Characterization Program

Miscellaneous

NCI CPTAC Assay Portal

Psychoneuroendocrinology. 2015 Jul;57:134-43. doi: 10.1016/j.psyneuen.2015.04.006. Epub 2015 Apr 14.

Sleep-deprivation reduces NK cell number and function mediated by β-adrenergic signalling

Beatriz H P De Lorenzo [1], Laís de Oliveira Marchioro [2], Carollina Ribeiro Greco [2], Deborah Suchecki [2]

Affiliations

PMID: 25929826 DOI: 10.1016/j.psyneuen.2015.04.006

Abstract

Reduction of sleep time triggers a stress response, leading to augmented levels of glucocorticoids and adrenaline. These hormones regulate components of the innate immune system such as natural killer (NK) and NKT cells. In the present study, we sought to investigate whether and how stress hormones could alter the population and function of NK and NKT cells of mice submitted to different lengths of paradoxical sleep deprivation (PSD, from 24 to 72 h). Results showed that 72h of PSD decreased not only NK and NKT cell counts, but also their cytotoxic activity against B16F10 melanoma cells in vitro. Propranolol treatment during PSD reversed these effects, indicating a major inhibitory role of beta-adrenergic receptors (β-AR) on NK cells function. Moreover, both corticosterone plasma levels and expression of beta 2-adrenergic receptors (β2-AR) in NK cells increased by 48 h of PSD. In vitro incubation of NK cells with dexamethasone augmented the level of β2-AR in the cell surface, suggesting that glucocorticoids could induce β2-AR expression. In summary, we propose that reduction of NK and NKT cell number and cytotoxic activity appears to be mediated by glucocorticoids-induced increased expression of β2-AR in these cells.

Keywords: Cytotoxicity; Murine melanoma; NK cells; Sleep deprivation; Stress; β(2)-Adrenergic receptor.

Related information

MedGen

LinkOut - more resources

Full Text Sources

ClinicalKey

Elsevier Science

Research Materials

NCI CPTC Antibody Characterization Program

Stress Medicine / Volume 7, Issue 1

Conference Proceeding

The impact of stressful life events on natural killer cells

Prof. Michael Schlesinger MD, Y. Yodfat MD

First published: January 1991
https://doi.org/10.1002/smi.2460070110
Citations: 12

Abstract

A study was carried out among members of a kibbutz to determine to what extent their capacity to cope with stressors affected natural killer (NK) cells and how this capacity was regulated by personal, family and social resources. The NK activity and NK cell markers were analysed among 92 kibbutz residents. A number of psychosocial parameters (including family function, social support, and demoralization) were assessed in parallel. A significant correlation was found between the capacity of individuals to cope with daily life stress and their cytotoxic NK activity. Individuals who were diagnosed as having anxiety neurosis had a significantly weaker NK activity and their population of Leu-11 positive cells was significantly lower than among those without such symptoms. No significant association could be determined between either NK cell activity or proportion of Leu-11 positive lymphocytes and any of the psychosocial parameters tested. Thus, while coping with stress has a significant effect on the NK system, further studies are required to elucidate the psychosocial mechanisms involved.

Grant S. Shields, George M. Slavich, Lifetime stress exposure and health: A review of contemporary assessment methods and biological mechanisms, Social and Personality Psychology Compass, 10.1111/spc3.12335, **11**, 8, (2017).
Wiley Online Library

Hideki Ohira, Social Support and Salivary Secretory Immunoglobulin a Response in Women to Stress of Making a Public Speech, Perceptual and Motor Skills, 10.2466/pms.98.3c.1241-1250, **98**, 3_suppl, (1241-1250), (2016).
Crossref

Donatella Marazziti, Fabio Ambrogi, Marianna Abelli, Elena Di Nasso, Mario Catena, Gabriele Massimetti, Marina Carlini, Liliana Dell'Osso, Lymphocyte subsets, cardiovascular measures and anxiety state before and after a professional examination, Stress, 10.1080/10253890601170563, **10**, 1, (93-99), (2009).
Crossref

HIDEKI OHIRA, SOCIAL SUPPORT AND SALIVARY SECRETORY IMMUNOGLOBULIN A RESPONSE IN WOMEN TO STRESS OF MAKING A PUBLIC SPEECH, Perceptual and Motor Skills, 10.2466/PMS.98.3.1241-1250, **98**, 3, (1241), (2004).
Crossref

Suzanne C. Segerstrom, Gregory E. Miller, Psychological Stress and the Human Immune System: A Meta-Analytic Study of 30 Years of Inquiry., Psychological Bulletin, 10.1037/0033-2909.130.4.601, **130**, 4, (601-630), (2004).
Crossref

Takao Miyazaki, Toshio Ishikawa, Hirofumi Iimori, Akiko Miki, Marcus Wenner, Isao Fukunishi, Noriyuki Kawamura, Relationship between perceived social support and immune function, Stress and Health, 10.1002/smi.950, **19**, 1, (3-7), (2003).
Wiley Online Library

Julie M. Turner-Cobb, Psychological and neuroendocrine correlates of disease progression, Neurobiology of the Immune System, 10.1016/S0074-7742(02)52016-X, (353-381), (2002).
Crossref

Eric P. Zorrilla, Lester Luborsky, James R. McKay, Robert Rosenthal, Arlene Houldin, Anne Tax, Ruth McCorkle, David A. Seligman, Kelly Schmidt, The Relationship of Depression and Stressors to Immunological Assays: A Meta-Analytic Review, Brain, Behavior, and Immunity, 10.1006/brbi.2000.0597, **15**, 3, (199-226), (2001).
Crossref

Rocco Domanico, Isiaah Crawford, Psychological Distress Among HIV-Impacted African-American and Latino Males, Journal of Prevention & Intervention in the Community, 10.1300/J005v19n01_04, **19**, 1, (55-78), (2000).

Crossref

Paola Borella, Annalisa Bargellini, Sergio Rovesti, Marina Pinelli, Roberto Vivoli, Valentina Solfrini, Gianfranco Vivoli, Emotional stability, anxiety, and natural killer activity under examination stress, Psychoneuroendocrinology, 10.1016/S0306-4530(99)00016-5, **24**, 6, (613-627), (1999).
Crossref

Loredana Fioroni, Maurizio Fava, Alessandro D. Genazzani, Fabio Facchinetti, Andrea R. Genazzani, Life events impact in patients with secondary amenorrhoea, Journal of Psychosomatic Research, 10.1016/0022-3999(94)90059-0, **38**, 6, (617-622), (1994).
Crossref

M. Schlesinger, Y. Yodfat, R. Rabinowitz, S. Bronner, J. D. Kark, Psychosocial Stress and NK Cells among Members of a Communal Settlement, Drugs of Abuse, Immunity, and AIDS, 10.1007/978-1-4615-2980-4_34, (247-254), (1993).
Crossref

ScienceDaily®

Your source for the latest research news

Sleep deprivation effect on the immune system mirrors physical stress

Date: July 1, 2012

Source: American Academy of Sleep Medicine

Summary: Severe sleep loss jolts the immune system into action, reflecting the same type of immediate response shown during exposure to stress, a new study reports. Researchers compared the white blood cell counts of 15 healthy young men under normal and severely sleep-deprived conditions. The greatest changes were seen in the white blood cells known as granulocytes, which showed a loss of day-night rhythmicity, along with increased numbers, particularly at night.

FULL STORY

evere sleep loss jolts the immune system into action, reflecting the same type of immediate response shown during exposure to stress, a new study reports.

Researchers in the Netherlands and United Kingdom compared the white blood cell counts of 15 healthy young men under normal and severely sleep-deprived conditions. The greatest changes were seen in the white blood cells known as granulocytes, which showed a loss of day-night rhythmicity, along with increased numbers, particularly at night.

"Future research will reveal the molecular mechanisms behind this immediate stress response and elucidate its role in the development of diseases associated with chronic sleep loss," said Katrin Ackermann, PhD, the study's lead author. "If confirmed with more data, this will have implications for clinical practice and for professions associated with long-term sleep loss, such as rotating shift work."

Previous studies have associated sleep restriction and sleep deprivation with the development of diseases like obesity, diabetes and hypertension. Others have shown that sleep helps sustain the functioning of the immune system, and that chronic sleep loss is a risk factor for immune system impairment.

For this study, white blood cells were categorized and measured from 15 young men following a strict schedule of eight hours of sleep every day for a week. The participants were exposed to at least 15 minutes of outdoor light within the first 90 minutes of waking and prohibited from using caffeine, alcohol or medication during the final three days. All of this was designed to stabilize their circadian clocks and minimize sleep deprivation before the intensive laboratory study.

White blood cell counts in a normal sleep/wake cycle were compared to the numbers produced during the second part of the experiment, in which blood samples were collected during 29 hours of continual wakefulness.

"The granulocytes reacted immediately to the physical stress of sleep loss and directly mirrored the body's stress response," said Ackermann, a postdoctoral researcher at the Eramus MC University Medical Center Rotterdam in the Netherlands.

The study, "Diurnal Rhythms in Blood Cell Populations and the Effect of Acute Sleep Deprivation in Healthy Young Men," was a collaborative effort between the Department of Forensic Molecular Biology at Erasmus MC University Medical Center Rotterdam and Chronobiology, Faculty of Health and Medical Sciences at the University of Surrey, United Kingdom. The laboratory study was conducted at the University of Surrey Clinical Research Centre.

Story Source:

Materials provided by **American Academy of Sleep Medicine**. *Note: Content may be edited for style and length.*

Journal Reference:

1. Katrin Ackermann, Victoria L. Revell, Oscar Lao, Elwin J. Rombouts, Debra J. Skene, Manfred Kayser. **Diurnal Rhythms in Blood Cell Populations and the Effect of Acute Sleep Deprivation in Healthy Young Men**. *SLEEP*, 2012; DOI: 10.5665/sleep.1954

Cite This Page:

RELATED STORIES

Losing a Night of Sleep May Increase Blood Levels of Alzheimer's Biomarker

Jan. 8, 2020 — A preliminary study has found that when young, healthy men were deprived of just one night of sleep, they had higher levels of tau -- a biomarker for Alzheimer's disease -- in their blood than when ...

Sleepless Nights Linked to High Blood Pressure

June 4, 2019 — A bad night's sleep may result in a spike in blood pressure that night and the following day, according to new research led by the University of Arizona. The study, to be published in Psychosomatic ...

Can't Sleep? Fruit Flies and Energy Drinks Offer New Clues

Nov. 8, 2018 — Like humans, fruit flies are active during the day, sleep at night and have similar sleep characteristics. A study has discovered a new gene and uncovered a mechanism that modulates sleep by ...

Sleep Duration Varies by Alcohol Drinking Patterns, Race, and Sex

June 13, 2016 — Compared to their white counterparts within each alcohol drinking pattern (never, moderate, excessive) investigated, black men and women were significantly more likely to get less than six hours of ...