Face masks to prevent transmission of respiratory diseases: Systematic review and meta-analysis of randomized controlled trials

Hanna M. Ollila\textsuperscript{1,2,3}, Markku Partinen\textsuperscript{4,5}, Jukka Koskela\textsuperscript{1,2,6}, Riikka Savolainen\textsuperscript{7}, Anna Rotkirch\textsuperscript{8}, and Liisa T. Laine\textsuperscript{9,10}

\textsuperscript{1}Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland
\textsuperscript{2}Broad Institute of MIT and Harvard, Cambridge, MA, USA
\textsuperscript{3}Stanford University School of Medicine, Palo Alto, CA, USA
\textsuperscript{4}Helsinki Sleep Clinic, Vitalmed Research Center
\textsuperscript{5}Department of Clinical Neurosciences, Clinicum, University of Helsinki, Helsinki, Finland
\textsuperscript{1}Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland
\textsuperscript{2}Broad Institute of MIT and Harvard, Cambridge, MA, USA
\textsuperscript{6}Helsinki University and Helsinki University Hospital, Clinic of Gastroenterology Helsinki, Finland
\textsuperscript{7}Newcastle University Business School, Newcastle-upon-Tyne, United Kingdom
\textsuperscript{8}Population Research Institute, Väestölitto – The Family Federation of Finland
\textsuperscript{9}The Wharton School of the University of Pennsylvania, Philadelphia, PA, USA
\textsuperscript{10}Department of Medical Ethics and Health Policy, The Perelman School of Medicine, Philadelphia, PA, USA

\textsuperscript{*}This version: December 2, 2020. Laine and Ollila gratefully acknowledge Academy of Finland for funding this research (Award number: 340551 LTL and 340539 HMO). Laine gratefully acknowledges funding from the National Institute on Aging of the National Institutes of Health under Award Number P30AG043073. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, nor of the Academy of Finland. Corresponding authors: Hanna M. Ollila, FIMM-EMBL Group leader, Institute for Molecular Medicine Finland, University of Helsinki, Finland; hanna.m.ollila@helsinki.fi and Liisa T. Laine, Department of Medical Ethics and Health Policy, The Perelman School of Medicine, Philadelphia, PA, USA; lainel@wharton.upenn.edu

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
ABSTRACT

Objective. To examine the effect of face mask intervention in respiratory infections across different exposure settings and age groups.

Design. Systematic review and meta-analysis.

Data sources. PubMed, Cochrane Central Register of Controlled Trials (CENTRAL), and Web of Science were searched for randomized controlled trials investigating the effect of face masks on respiratory infections published by November 18th 2020. Our reporting follows the PRISMA guidelines.

Eligibility criteria for selecting studies. Randomized controlled trials investigating the effect of face masks in respiratory infections and influenza-like illness across different exposure settings and age groups. Two reviewers independently performed the search, extracted the data, and assessed the risk of bias. A random effects meta-analysis with risk ratio, risk difference, and number needed to treat were performed. Findings in exposure settings, age groups, and role of non-compliance were examined using a subgroup analysis.

Results. Total of 17 studies were included, with N = 11,601 individuals in intervention and N = 10,286 in the control group with follow-up duration from 4 days to 19 months). 14 trials included adults (and children) and 3 included children only. 12 studies suffered from non-compliance in the treatment arm and 11 in the control arm. All studies were intent-to-treat analyses, and, thus, non-compliance can bias individual intent-to-treat estimates towards zero. Four out of seventeen studies supported use of face masks. A meta-analysis of all 17 studies found no association between face mask intervention and respiratory infections (RR = 0.9046 [0.777 - 1.053], p = 0.196, p fixed effect = 0.0006). However, a meta-analysis using odds ratios adjusted for age, sex, and vaccination (when available) suggests protective effect of the face mask intervention (17 studies, OR = 0.850 [0.736 - 0.982], p=0.027). A subgroup meta-analysis among adults with (unadjusted) risk ratios found a decrease in respiratory infections (14 studies, RR = 0.859 [0.750 - 0.983], p = 0.026, and 4 studies with a combined face masks and hand hygiene intervention RR = 0.782 [0.696 - 0.879], p < 0.0001). Finally, the face mask use is also supported by a meta-regression adjusting the effect estimates for non-compliance in the controls (17 studies RR = 0.87 [0.780 - 0.980], p = 0.017).

Conclusion. Our findings support the use of face masks to prevent respiratory infections.
1 Introduction

The COVID-19 and other pandemics are a scourge causing severe losses on health, economy, and well-being [1, 2]. COVID-19, caused by SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2), can spread through droplet-mediated transmission through contaminated surfaces and air [3–6]. Non-pharmaceutical interventions (NPIs), such as maintaining physical distance, appropriate hand hygiene, and face masks have been adopted as the primary tools to limit the number of COVID-19 infections [7] while the vaccines are being developed and pharmaceuticals are studied for repurposing.

Prior to COVID-19 pandemic, the use of face masks by the general public was considered a relatively new policy tool in preventing person-to-person transmission on a global scale. Face masks are widely used in health care settings. Prediction models suggest that universal use of face masks in public may have a substantial preventive impact on disease spread, even without medical masks or 100% compliance [5,10]. In addition, a pooled meta-analysis of the spread of infectious viral diseases of up to 172 studies showed a consistent effect regarding the efficacy of face masks in preventing infections by SARS-CoV-2 and the betacoronaviruses that cause severe acute respiratory syndrome, and Middle East respiratory syndrome [7].

However, the most robust type of evidence on the efficacy of face masks use among the general public – that from randomized controlled trials – has been noted as being only suggestive. [11] For example, many of the randomized controlled trials have documented non-compliance either in the face mask intervention group [12–23] or in the control group [12, 15, 17, 18, 21, 23–28]. Because these studies estimate the intention-to-treat effect of face masks, non-compliance can underestimate the magnitude of the treatment effect of face masks use for a given randomized controlled trial.

The aim of this systematic review and meta-analysis was to examine the evidence from randomized controlled trials of face masks in the context of COVID-19 or diseases which spread through similar mechanisms to COVID-19: respiratory infections. An earlier systematic review and meta-analysis has investigated the effect of face masks by focusing on the use of cloth masks [29] in non-health care settings while [30] combined various types of studies, including RCTs, case-control studies and cohort studies. Our review complements these studies by focusing solely on randomized controlled trials in different exposure settings (hospital, household, and community) and age groups (adults vs. children). Moreover, we study the role of non-compliance in treatment and control groups and whether the results differ if the face mask use intervention included hand hygiene.
guidance or not.

2 METHODS

This systematic review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses. Our review protocol was registered on PROSPERO in November 2020 (registration number CRD42020205523).

2.1 SEARCH STRATEGY

We performed the searches using the Cochrane Central Register of Controlled Trials (CENTRAL), PubMed, and Web of Science (science and social science citation index). We performed the PubMed search using Medical Subject Headings (MeSH) listed in Supplement A. In other search engines, we used the following search terms: facemasks/face masks AND/OR infection. The full search protocol with the criteria are described in Supplement A. We limited the searches to randomized controlled trials on humans published by November 18th 2020. We did not limit the searches by language. The search results were uploaded on Endnote, and the unique citations were kept and screened.

2.2 STUDY SELECTION, INCLUSION AND EXCLUSION CRITERIA

We included randomized controlled trials on humans (general population and health care personnel in a risk of contracting respiratory infectious diseases) that compared face mask use (FFP1, FFP2, FFP3, cloth mask or surgical mask) or face mask and hand hygiene or face mask and education with no face mask use. We did not make exclusions based on a setting, instead, we included interventions that were executed in various settings, such as in health care, community, or household. Our included measure was the relative risk for infection.

We excluded interventions that compared different types of face masks to each other (in which the comparison group were assigned to use a face mask). We did not exclude any studies based on age and gender or have exclusion criteria based on sample sizes or follow-up periods. We included all the studies with a whole text available (including pre-prints) while we excluded the studies which had only an abstract available. Table A1 in Supplement A provides a detailed summary of the inclusion and exclusion criteria.

Two authors (HMO and LTL) executed the search. The authors (HMO and LTL) independently
reviewed all the titles and abstracts to define the papers that could potentially be included in the systematic review. After this, both authors independently screened the articles and determined whether they met the inclusion and exclusion criteria. The disagreements between the two authors were resolved by discussion.

2.3 Data extraction

Two authors (HMO and LTL) independently extracted the data which included (1) study setting (time, country, population); (2) intervention details (randomization level, follow-up, type of mask, other interventions, case or index case definition); (3) outcome measures (effect size or N per group); (4) compliance measure; and (5) study results for the effects of face mask use. Two other reviewers (JK and RS) checked the extracted data for errors.

2.4 Risk-of-bias assessment

Two review authors (HMO and LTL) independently assessed the risk of bias using the Cochrane Risk of Bias tool. Any discrepancies or unusual patterns were resolved by consensus. The following characteristics were evaluated: Random sequence generation, allocation concealment (selection bias), blinding of participants and personnel (performance bias), blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), selective reporting (reporting bias), and non-compliance in the treatment group and control group. The risks were categorized as low risk, unclear risk or high risk of bias. Following the Cochrane tool for risk assessment, we denoted the overall risk of bias as low if all the categories were at a low risk of bias, high if at least one domain was at a high risk of bias. We denoted the overall risk of bias as unclear if at least one domain was at an unclear risk of bias and no domain was at a high risk.

2.5 Data analysis

The results for all the outcomes were expressed as risk ratios while we used 95% confidence intervals. We combined the estimates using a random-effects meta-analysis, based on the assumption that the existence of methodological and clinical heterogeneity potentially affecting the results was likely. We estimated the between-study variance by using the DerSimonian and Laird method of moments estimator. We calculated the 95 percent confidence intervals using the Wald method.

Taking the overall risk into account is important because it helps in avoiding the bias caused by prioritizing one category over others as any source of bias can be problematic.\[32\]
We assessed heterogeneity and quantified statistical inconsistency by using the $\xi^2$ test and the $I^2$ statistic, respectively. We used stratified meta-analyses to explore heterogeneity in the effect estimates according to: non-compliance, study populations, and settings. We studied how non-compliance in controls (using a face mask) is associated with the intervention effects in the meta-analysis with a meta-regression.

The small study effects were studied by generating contour-enhanced funnel plots to examine the bias in the results of the meta-analysis (the tendency for intervention effects from smaller studies to differ from those estimated in larger ones, which can result from reporting biases, methodological or clinical heterogeneity or other factors).

We conducted all the analyses using the meta, metafor and dmetar packages in R version 4.02 and meta package in Stata version 16.

3 Results

3.1 Search results and study characteristics

Our search resulted in 2,354 unique publications. After the review, we retained 17 randomized controlled trials of face mask use while 2,337 articles were excluded because they did not meet the inclusion criteria (Figure 1). Altogether, these studies included 11,601 study participants in the treatment group and 10,286 in the control group (Table 1). The duration of follow-up varied from 4 days to 19 months. The studies included a variety of environmental settings: pilgrims (3 studies), college students (2 studies), healthcare (4 studies) to household environment (7 studies). Six trials were performed in a community setting [12, 16, 24]. Three included children only [20, 22, 23] and 14 trials included to both adults and children [12, 19, 21, 24, 28].

3.2 Characteristics of included studies

Table 1 summarizes the characteristics of each study. The trials were carried out in eleven countries in several continents: Australia [20], China [21, 27], Denmark [16], France [25], Germany [22], Hong Kong [17, 18], Japan [26], Saudi Arabia [12, 15, 24], Thailand [23], the United States [13, 14, 19], and Vietnam [28].
Total 2,430 records identified through database search
4 additional records identified through other sources
Records after duplicates removed (n = 2,354)
Records screened (n = 2,354)
Records excluded (n = 2,327)
Full-text articles assessed for eligibility (n = 27)
Studies included in qualitative synthesis (n = 17)
Studies included in quantitative synthesis (meta-analysis) (n = 17)

Full-text articles excluded, with reasons
(2 = duplicate with different publication source
1 = was a review
1 = did not provide numbers for intention to treat analysis
1 = used existing data from Cowling et al., 2009
1 = had no control arm
1 = overlapping data with another study
1 = respirator vs. medical mask
1 = studied tolerability of masks
1 = measures surgical infection and not respiratory infection)

Figure 1: PRISMA flow diagram of included articles.
**Randomization**

2006 - 9 days + one

2007 - Population

1 month

**Disease and/or index case definition**

446 no intervention

**Type of mask and other interventions**

- Mask group was 2994 in control group
- Controls; 10 tents; 89 in control group
- Adults over 18 years
- Facemask N = 148
- Controls N = 158

and 306 contacts

**Participants**

- Total N = 995
- Students
- Verbal and printed instructions, mask, Cat No: 1816
- Pump)

- Medical procedure masks with appropriate hand sanitizer and mask use
- Control with health education: health leaflet
- Written materials detailing education
- Control with health education: health leaflet
- Laboratory-confirmed influenza in a subset
- Facemask and hand hygiene intervention:

**Mask only**

- Use of 3.7 hours a day.
- 4 days with an average use of 2 masks per day and a duration of

**Compliance**

**Symptoms of ILI**

42.4% of participants:

- Throat swab specimens were tested for influenza
- Laboratory-confirmed influenza in a subset
- Asymptomatic cases

**Intervention**

- Throat swab specimens were tested for influenza
- Laboratory-confirmed influenza in a subset
- Facemask and hand hygiene intervention:

**Inclusion**

- Index case had facemask
- No mask in controls
- Surgery masks with earloops, 3 plys, anti fog.
- Reported being outside the home among
- 98%; made in China)
- Aged 18 years or older
- - 306 contacts
- - Control; 10 tents; 89 in control group
- - Adults over 18 years
- - Facemask group
- - Control group reported use of 1.51 times per day.
- - Mask only group reported an average use of hand sanitizer
- - Face mask and hand hygiene group wore their mask, on

**Follow-up**

- Mask use compliance
- Control group reported use of 1.51 times per day.
- Mask only group reported an average use of hand sanitizer
- Average number of mask hours per day

**Reported**

- Use of 3.7 hours a day.
- 4 days with an average use of 2 masks per day and a duration of
Absence of randomization in the no mask arm. Wearing was not routine during the study period.

The rates of all outcomes were shown no difference between arms. No association in the total analysis with facemask preventing respiratory clinical infection or with and without fit testing. The secondary attack ratios were no significant differences between intervention arms.

Compliance with the hand hygiene intervention seemed low. Over 25% of the index cases in the control and hand hygiene group subjects had no hand hygiene education on healthy diet and exercise in the preceding 14 days. None of whom had reported acute respiratory illness in the preceding 14 days. Persons who reported at least 2 symptoms of ILI were residents aged at least 2 years.

Over 25% household contacts in the face mask group did not wear any mask while on hospital property unless required to do so as part of their job duties. Over 25% of the index cases in the control and hand hygiene groups reported higher adherence to the intervention compared with The intervention group compared to The intervention group compared to 50% of index patients in the facemask plus hand hygiene group.

Adherence to the hand hygiene intervention seemed low. Subjects in the no mask group refrained from wearing a face mask while on hospital property unless required to do so as part of their job duties. Subjects in the no mask group refrained from wearing a face mask while on hospital property unless required to do so as part of their job duties. Over 25% of the index cases in the control and hand hygiene group subjects had no hand hygiene education on healthy diet and exercise in the preceding 14 days. None of whom had reported acute respiratory illness in the preceding 14 days. Persons who reported at least 2 symptoms of ILI were residents aged at least 2 years.

Over 25% household contacts in the face mask group did not wear any mask while on hospital property unless required to do so as part of their job duties. Over 25% of the index cases in the control and hand hygiene groups reported higher adherence to the intervention compared with The intervention group compared to The intervention group compared to 50% of index patients in the facemask plus hand hygiene group.

Adherence to the hand hygiene intervention seemed low. Subjects in the no mask group refrained from wearing a face mask while on hospital property unless required to do so as part of their job duties. Subjects in the no mask group refrained from wearing a face mask while on hospital property unless required to do so as part of their job duties. Over 25% of the index cases in the control and hand hygiene group subjects had no hand hygiene education on healthy diet and exercise in the preceding 14 days. None of whom had reported acute respiratory illness in the preceding 14 days. Persons who reported at least 2 symptoms of ILI were residents aged at least 2 years.

Over 25% household contacts in the face mask group did not wear any mask while on hospital property unless required to do so as part of their job duties. Over 25% of the index cases in the control and hand hygiene groups reported higher adherence to the intervention compared with The intervention group compared to The intervention group compared to 50% of index patients in the facemask plus hand hygiene group.

Adherence to the hand hygiene intervention seemed low. Subjects in the no mask group refrained from wearing a face mask while on hospital property unless required to do so as part of their job duties. Subjects in the no mask group refrained from wearing a face mask while on hospital property unless required to do so as part of their job duties. Over 25% of the index cases in the control and hand hygiene group subjects had no hand hygiene education on healthy diet and exercise in the preceding 14 days. None of whom had reported acute respiratory illness in the preceding 14 days. Persons who reported at least 2 symptoms of ILI were residents aged at least 2 years.

Over 25% household contacts in the face mask group did not wear any mask while on hospital property unless required to do so as part of their job duties. Over 25% of the index cases in the control and hand hygiene groups reported higher adherence to the intervention compared with The intervention group compared to The intervention group compared to 50% of index patients in the facemask plus hand hygiene group.
MacIntyre et al., 2015

<table>
<thead>
<tr>
<th>2015</th>
<th>Healthcare workers</th>
<th>Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>1607 healthcare workers in 15 hospitals (control group)</td>
<td>36-62 (59±9)</td>
</tr>
<tr>
<td>2015</td>
<td>Medical mask (KN95)</td>
<td>Total N = 925</td>
</tr>
<tr>
<td>2015</td>
<td>Medical mask with a mask</td>
<td>Total N = 925</td>
</tr>
<tr>
<td>2015</td>
<td>Handwashing kit</td>
<td>Total N = 925</td>
</tr>
<tr>
<td>2015</td>
<td>Surgical mask</td>
<td>Total N = 925</td>
</tr>
<tr>
<td>2015</td>
<td>Medical mask</td>
<td>Total N = 925</td>
</tr>
<tr>
<td>2015</td>
<td>Cloth mask</td>
<td>Total N = 925</td>
</tr>
<tr>
<td>2015</td>
<td>No mask</td>
<td>Total N = 925</td>
</tr>
<tr>
<td>2015</td>
<td>18% of the control group</td>
<td>Total N = 925</td>
</tr>
<tr>
<td>2015</td>
<td>The rate of ILI was lower in the mask arm compared to the control group</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Descriptive table outlining study characteristics in this systematic review and meta-analysis.
<table>
<thead>
<tr>
<th>Abdin et al., 2005</th>
<th>Unclear risk</th>
<th>Unclear risk</th>
<th>Unclear risk</th>
<th>Unclear risk</th>
<th>Low risk</th>
<th>Low risk</th>
<th>Unclear risk</th>
<th>Unclear risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aiello et al., 2010</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
</tr>
<tr>
<td>Aiello et al., 2012</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
</tr>
<tr>
<td>Alfellai et al., 2020</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>High risk</td>
</tr>
<tr>
<td>Bundgaard et al., 2020</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
</tr>
<tr>
<td>Barasheed et al., 2014</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>Canini et al., 2010</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>Cowling et al., 2008</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
</tr>
<tr>
<td>Cowling et al., 2009</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
</tr>
<tr>
<td>Jacobs et al., 2009</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
</tr>
<tr>
<td>Larson et al., 2010</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
</tr>
<tr>
<td>Macintyre et al., 2009</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
</tr>
<tr>
<td>Macintyre et al., 2011</td>
<td>High risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>Macintyre et al., 2015</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
</tr>
<tr>
<td>Macintyre et al., 2016</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
</tr>
<tr>
<td>Suess et al., 2012</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>High risk</td>
</tr>
<tr>
<td>Simmerman et al., 2011</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>Bundgaard et al., 2020</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear risk</td>
</tr>
</tbody>
</table>

**Figure 2:** Bias assessment. Non-compliance in the treatment or control group: high risk if the reported non-compliance was greater than 50%, unclear if between 30-50%, and low if under 30%.

**Figure 3:** Review authors’ judgments on each risk-of-bias item as percentages across all the included studies.
3.3 Assessment of intervention: face mask use

In addition to conducting interventions in diverse settings (community, hospital, household) and age groups (adults, children), the interventions themselves varied. In some of the interventions, the treatment group received an education leaflet in addition to face masks [12], while, in others, the intervention included a weekly supply of face masks and a plastic bag for storage and daily disposal [15]. The type of face mask varied from cloth masks to medical masks with ear loops. In some studies, the level of information on the face mask use was unclear. Some trials had a specific hand hygiene and face mask arm [13, 14, 18, 19] in which the treatment group intervention included also a hand sanitizer.

Four studies [13, 14, 19, 24] found a protective effect of face masks in the intention-to-treat analysis. Two of these studies had a follow-up length of 6 weeks and one up to 19 months, which were the longest follow-up times among the 17 included interventions. In addition, two additional studies showed an association in the per protocol analysis [18, 34]. These studies were source protection studies where an early intervention within 36h was associated with a reduced number of respiratory infections in the contacts, suggesting that face masks may be most efficient if adopted early on during the exposure.

3.4 Risk of bias across the studies

Figure 2 summarizes the risk of bias on the study level. The observed bias was low or unclear in the majority of the 17 randomized controlled trials. In the instances in which a bias was found, the main concerns were related to non-compliance either in the treatment (12 studies [12–23]) or in the control arm so that treated individuals did not use the mask while individuals in the control arm did use it (11 studies [12, 15, 17, 18, 21, 23–28]). Almost all the trials had an increased a risk of bias due to unclear or a lack of blinding. Obviously, blinding per mask use is challenging due to the visible nature of face mask. In addition, one study could not allocate the control arm randomly due to local health regulations, so it recruited a separate (non-randomized) control group and examined primarily differences between face masks [27]. There were some concerns due to the lack of blinding at the stage of identification of symptoms per treatment arm (13 studies [12, 19, 21, 24, 26, 28]). Similarly, it is unclear if the outcome assessment was blinded in any of the 17 studies (Figure 2). Details about random sequence generation and allocation concealment were unclear for some trials. A summary of the proportion of the trials that were at low, unclear, and high bias for each domain
is shown in (Figure 3). We found no evidence of a publication bias by a visual examination of funnel plots (Figure 7) or by an analysis based on Egger’s tests: $\beta = -0.08$, $se = 0.52$, $p = 0.88$.

3.5 FACE MASKS AND RESPIRATORY INFECTIONS

In total, there were 1,330 events among treatment arms (N total face mask = 11,601) and 1,335 events in control arms (N total controls = 10,286). A random effects meta-analysis of the results showed that, at the end of the trials. The median follow-up time was 9 days and ranged from 4 days to 19 months. There was no association in the meta-analysis across all 17 studies ($RR = 0.9046 [0.777 - 1.053]$, $p = 0.196$, $p$ fixed effect = 0.0006). However, there was a statistically significant reduction in respiratory infections in the individuals over 15 years of age ($RR = 0.859 [0.750 - 0.983]$, $p = 0.027$ Figure 4), corresponding to a risk difference of $-0.016 [-0.0320; -0.0002]$. These effects were relatively small, and significant between-study heterogeneity within this population remained ($\tau^2 = 0.023$, $I^2 = 53.0\%$ [13.7\% - 74.4\%], $p = 0.01$).

Among the studies, the adjusted values were also provided in a subset of studies. In some studies, these values differed substantially from the unadjusted ones [23] or, as mentioned above, there was non-compliance in the treatment or control arm or both. For this reason, we performed a secondary analysis using adjusted odds ratios in the original articles when available. In this analysis, the overall effect in all the 17 RCTs - including in children - showed a protective effect ($OR = 0.850 [0.736 - 0.982]$, $p = 0.0269$ Figure A1).

It is possible that individual studies bias the estimate. Therefore, we performed a meta-analysis through a leave-one-out analysis to examine if a systematic association to a given direction was observed after excluding a given study. The effect sizes were systematically at $RR < 1$ with all the analyses. However, one of the largest studies [15] had a significant level of non-compliance in the treatment group with 49% of the controls using face masks. As shown in Figure 8, this study biases the association towards the null hypothesis. Similarly, we estimated significance of non-compliance on the finding. A subgroup meta-analysis without the studies with non-compliance of over 10% in the control group showed a protective effect of face masks ($0.776 [0.717 - 0.841]$, $p < 0.0001$). Similarly, including those studies in which controls used face masks for over 10% of the time, we observed a statistically significant difference between the studies where controls did not use masks vs. where controls used masks (meta-regression $p=0.011$). Similarly, adjusting the RR estimate with a meta-regression in all the 17 RCT studies suggests a statistically significant association (RR adjusted for compliance = 0.87[0.78-0.98], $p=0.017$), indicating that non-compliance has weakened...
Environments differ by their risk of contracting respiratory viruses due to having varying amount of viral particles or a different length of exposure. As a result, effects of face masks likely differ by the length and the setting of the exposure. We investigated potential differences by conducting a subgroup meta-analysis of different environments: community and hospital or household settings by focusing on studies that included adults (Figure 5). In the random effects meta-analysis with the raw reported number of individuals, the result suggests an effect in the hospital or household settings \((RR = 0.803 \ [0.727 - 0.887], p < 0.0001)\), while the effect was similar in the community settings although statistically insignificant \((RR = 0.838 \ [0.689 - 1.012], p = 0.077, Figure 5)\). It is possible that these large confidence intervals in the community setting result from non-compliance (between 10% and 50% of the control arms using masks) in three out of the six studies that assessed community transmission and from relatively low compliance in the treatment arms. In the twelve of the trials including adults, the intervention consisted solely of face mask use while in four the intervention included also guidance on appropriate hand hygiene together with the face mask use.
The subgroup analysis for the face mask with hand hygiene guidance resulted in the strongest protection ($RR = 0.785 \ [0.695 - 0.886]$, $p < 0.0001$), Figure 6. In contrast, one large study where the controls used face masks 49% of the time biased the estimate towards no effect in the face mask only group ($RR = 0.93 \ [0.839 - 1.039]$, $p = 0.183$) included while, after exclusion, $RR = 0.868 \ [0.769 - 0.979]$ $p = 0.021$. 

**Figure 5**: Random effects meta-analysis of the relative risk of respiratory infections, subgroup analysis by setting (community setting vs. hospital or household). The figure includes both fixed-effects and random-effect models.
**Figure 6**: Random effects meta-analysis of the relative risk of respiratory infections, subgroup analysis taking other NPI into account (mask and hand hygiene vs. mask only). The figure includes both fixed-effects and random-effect models.
**Figure 7:** Contour enhanced funnel plot for the random effects meta-analysis of the log risk ratio of face mask use on respiratory infections.

**Figure 8:** Robustness - leave one out analysis.
3.6 Number needed to treat

We approximate the effect of masks on population health by exploring the number needed to treat, that is, how many individuals need to wear a mask to prevent one person from contracting a respiratory infection. The number needed to treat depends on such infections at the population level. If there are few infections, a larger number of mask users will be needed to prevent one infection. Based on the results from this meta-analysis and assuming a low baseline risk of 0.01, NNT is 455. With a larger baseline risk, 0.05, NNT is 91, and with a higher still risk, 0.2, NNT becomes 23. Similarly, the risk difference ranged from −3% to 0% in adults ($RD = −0.016 \quad [−0.0320 − −0.0002], \quad p = 0.048,$) and was larger with appropriate hand hygiene $RD = −0.051 \quad [−0.116 − −0.014], \quad p = 0.12$).

4 Discussion

4.1 Main findings

This systematic review and meta-analysis of 17 RCT studies examined whether face masks prevent respiratory infections. 4 out of 17 studies supported the use of masks in the intention-to-treat analysis. In addition, the analysis of the adjusted ORs showed a protective effect of face masks. Similarly, once the studies with non-compliance in the control groups were removed, the corresponding meta-analysis showed a protective effect of face masks. Finally, adjusting the effect estimates through a meta-regression supported the use of masks across all the 17 studies. The association was seen both with mask use alone and when masks were combined with appropriate hand hygiene. This result is aligned with the current evidence that NPIs are most efficient when used together. It is worth noting that, despite the relatively large between-study heterogeneity and small effect sizes in the individual studies, the findings did support use of face masks. Therefore, these findings together with the mounting other evidence suggest that face masks may be considered as a useful NPI for respiratory infections, including COVID-19.

4.2 Quality of evidence

We found 17 randomized controlled trials that had assessed whether masks affect the number of respiratory infections. Other earlier studies have been conducted using case-control settings or with masks with a strong filtering capacity [7]. Earlier systematic reviews and meta-analyses have investigated, for example, the effect of face masks by combining types of studies, including RCTs,
case-control studies and cohort studies \cite{30} or cloth masks \cite{29} in non-health care settings. The findings from our systematic review and meta-analysis is in line with the conclusions from these earlier meta-analyses conducted in different settings. In contrast, by including as a full set of studies as possible, we are better powered in estimating the effect of a mask intervention.

While the intention-to-treat analysis yields an unbiased estimate of the efficacy of the face mask intervention, its magnitude is biased downwards relative to the actual treatment effect of face masks. While the overall quality of the earlier trials is solid, there were biasing factors across the studies, including a compliance bias either because of low compliance in the mask arm or use of face masks in the control arm. According to the sensitivity analysis, these findings may bias the estimates towards the null hypothesis.

In addition, as the effects with hand hygiene seems to be stronger than with mask use alone, someone might conclude that hand hygiene is driving the association while mere face masks do not protect from respiratory infections. Indeed, while masks have been shown to be effective in themselves, their impact and, therefore, efficacy is largest when combined with other protective measures \cite{7}. Also in our study, the effect of masks was further accentuated when combined with complementary measures, such as improved hand hygiene \cite{13, 14}. Furthermore, other complementary measures for disease control, such as physical distancing measures, have an impact on the spread of diseases and the number of particles in the air and, hence, also add to the effect of face masks.

Indeed, in a review \cite{35}, the estimated number needed to mask to prevent one infection ranged from three (N95 masks) to six (face masks), and the number is higher still when the infection risk is low to start with. Clearly, these NNTs are only approximations since the reproduction number R differs between viral infections. Similarly, if there are no active infections, the NNT will be infinite: no infections can be prevented as none are present in the population.

With these limitations in mind, we calculated that, for respiratory infections, the NNT might range from 23 to 455. To put this into context, let us presume that, in a metropolitan area with a population of one million, 30% of the residents use face masks. With NNT=455, this might prevent 600 respiratory infections. This effect size is comparable to the NNT of pharmaceuticals. For example, the NNT for statin, one of the most widely prescribed drugs, in primary prevention of major vascular events at low levels of a CVD risk (5-10% within 5 years) ranges from 67 to 170 and is of a similar scale to face masks \cite{36}.

We show that the studies where hand hygiene was assessed together with mask use, effects with
multiplicative protective measures were seen. Our results support use of face masks in preventing respiratory infections and, hence, the WHO guidelines that recommend the use of face masks together with physical distancing and hand hygiene in controlling the spread of COVID-19.

4.3 LIMITATIONS

First of all, the population studied here had residual heterogeneity. Indeed, as respiratory infections are time- and exposure-dependent, it is possible that differences in follow-up times and in symptom assessments (ILI, respiratory illness or COVID-19) have affected the power to detect associations. Second, while all the studies reported the numbers in the treatment and control arms, we did not have access to raw data and could not adjust the analysis by within-study variables. As a work-around, we performed a meta-analysis including within-study adjusted odds ratios. However, this method comes with limitations of its own as because, in practice, no studies have exactly the same covariate definitions, which bias the estimates. Third, the mask types and instructions for mask use together with follow-up times varied by study, which likely increases between-study heterogeneity. We accounted for the biases through subgroup analysis by age group, setting and non-compliance in controls and meta-regression by non-compliance in controls.

4.4 CONCLUSIONS AND FUTURE IMPLICATIONS

Our meta-analysis using 17 randomized controlled trials across different exposure settings and age groups provides support for the public health policy of face mask use to limit the spread of infectious respiratory diseases, such as COVID-19. Our analysis suggests that face masks can decrease the probability of contracting a respiratory infection and can be particularly useful when combined with other personal protection methods.

Recommendations and clear communication about the benefits of face masks should be given by policymakers to limit the number of respiratory infections and, ultimately, deaths in respiratory disease pandemics, thus providing time for vaccine and treatment development.

CONFLICT OF INTEREST

The authors declare no conflict of interest.
Author contribution

HMO and LTL conceived the study and conducted the main analysis. JK, MP, RS and AR assisted with the analyses and drafted the initial manuscript with HMO and LTL. All the authors participated in the interpretation, contributed to the drafts of the manuscript, and approved the final version. HMO and LTL are the guarantors and ensure that all the listed authors meet the authorship criteria and that no others meeting the criteria have been omitted.

Patient and public involvement

The current research was not informed by patient and public involvement because it used secondary data. Patients or the public were not involved in the design, conduct, reporting or dissemination plans of our research.
REFERENCES


Supplement

A Study criteria

Table A1: Study criteria

2. No setting restrictions were made (included all settings health care, community, and households).
3. Intervention type: face mask (FFP1, FFP2, FFP3, cloth mask or surgical mask).
4. Comparison group: no face mask. Exclusion of studied with comparison using a different type of face mask.
5. Mask use could occur with or without hand hygiene or other measures (e.g. use education).
6. Outcomes: Relative risk for infection, safety and efficacy, slowing of infection at the population, health care setting, or household level.
6. Included: All aged and gender.
7. Publication format. Whole text available, preprints included. Exclusion criteria; only abstract available.
8. Language of original publication; Primary search in English. No exclusion for other languages.
9. Sample size and follow up did not have exclusion criteria.
11. Excluded: Studies which studied face masks to examine effect on empathy or non-respiratory illness (eg. wound infection).

B Search terms

Our literature search can be replicated using using the following protocol.

B.1 Cochrane search

Keywords — facemasks, infection OR “face masks”. Search with i) ”facemasks, infection” resulted 47 items. Search with ii) ”face masks, infection” resulted 146 items.

- Take those results found on RCT tab (add to table for flow chart).
- Compute total number from all the searches before duplicates (add to table for flow chart)
- Remove duplicates and record their number (add to table for flow chart)
- Keep only those RCT studies that measure the respiratory infection of the person wearing the mask OR those that measure protection from respiratory infections towards others.
- Do not keep articles that measure empathy or other related traits

B.2 Pubmed search

MeSH terms — On tab Search trial (”Masks”[Mesh] OR ”Respiratory Protective Devices”[Mesh] OR ”mask” OR ”facial mask”) AND (”infection”OR ”Encephalitis, Viral”[Mesh] OR ”Viral Load”[Mesh]

24
iv) Search resulted in 2,161 items (including duplicates).

- Based on abstracts of the studies from Pubmed.
- Take RCTs.
  - Keep only those RCTs that measure the respiratory infection of the person wearing the mask OR those that measure protection from respiratory infections towards others.
  - Do not keep articles that measure empathy or other related traits.

B.3 Web of Sciences

Search terms — Facemask AND infection AND randomized controlled trials, Face mask AND infection AND randomized controlled trials

v) Search with "Facemask AND infection AND randomized controlled trials" resulted in 63 items.

vi) Search with "Face mask AND infection AND randomized controlled trials" resulted in 13 items.

C Supplementary materials
### Figure A1: Random effects meta-analysis of the adjusted odds ratio risks of respiratory infection.

The figure includes both fixed-effects and random-effect models.

<table>
<thead>
<tr>
<th>Study</th>
<th>TE</th>
<th>seTE</th>
<th>Odds Ratio</th>
<th>OR</th>
<th>95%-CI</th>
<th>Weight (fixed)</th>
<th>Weight (random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MacIntyre et al., 2016</td>
<td>-1.13</td>
<td>1.1611</td>
<td></td>
<td>0.32</td>
<td>[0.03; 3.14]</td>
<td>0.1%</td>
<td>0.4%</td>
</tr>
<tr>
<td>Barasheed et al., 2014</td>
<td>-0.93</td>
<td>0.4550</td>
<td></td>
<td>0.39</td>
<td>[0.16; 0.96]</td>
<td>0.9%</td>
<td>2.3%</td>
</tr>
<tr>
<td>MacIntyre et al., 2011</td>
<td>-0.72</td>
<td>0.7100</td>
<td></td>
<td>0.49</td>
<td>[0.12; 1.95]</td>
<td>0.4%</td>
<td>1.0%</td>
</tr>
<tr>
<td>Suess et al., 2012</td>
<td>-0.67</td>
<td>0.4735</td>
<td></td>
<td>0.51</td>
<td>[0.20; 1.29]</td>
<td>0.9%</td>
<td>2.1%</td>
</tr>
<tr>
<td>Aiello et al., 2010</td>
<td>-0.50</td>
<td>0.1258</td>
<td></td>
<td>0.60</td>
<td>[0.47; 0.77]</td>
<td>12.1%</td>
<td>12.5%</td>
</tr>
<tr>
<td>Aiello et al., 2012</td>
<td>-0.32</td>
<td>0.1930</td>
<td></td>
<td>0.72</td>
<td>[0.50; 1.06]</td>
<td>5.1%</td>
<td>8.4%</td>
</tr>
<tr>
<td>Larson et al., 2010</td>
<td>-0.20</td>
<td>0.0832</td>
<td></td>
<td>0.82</td>
<td>[0.70; 0.97]</td>
<td>27.6%</td>
<td>15.8%</td>
</tr>
<tr>
<td>Bundgaard et al., 2020</td>
<td>-0.20</td>
<td>0.2100</td>
<td></td>
<td>0.82</td>
<td>[0.54; 1.24]</td>
<td>4.3%</td>
<td>7.6%</td>
</tr>
<tr>
<td>Cowling et al., 2009</td>
<td>-0.15</td>
<td>0.2957</td>
<td></td>
<td>0.86</td>
<td>[0.48; 1.54]</td>
<td>2.2%</td>
<td>4.7%</td>
</tr>
<tr>
<td>Jacobs et al., 2009</td>
<td>-0.14</td>
<td>1.4604</td>
<td></td>
<td>0.87</td>
<td>[0.05; 15.23]</td>
<td>0.1%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Cowling et al., 2008</td>
<td>-0.13</td>
<td>0.4843</td>
<td></td>
<td>0.88</td>
<td>[0.34; 2.27]</td>
<td>0.8%</td>
<td>2.1%</td>
</tr>
<tr>
<td>Abdin et al., 2005</td>
<td>-0.04</td>
<td>0.1451</td>
<td></td>
<td>0.96</td>
<td>[0.73; 1.28]</td>
<td>9.1%</td>
<td>11.2%</td>
</tr>
<tr>
<td>Canini et al., 2010</td>
<td>-0.01</td>
<td>0.3395</td>
<td></td>
<td>0.99</td>
<td>[0.51; 1.93]</td>
<td>1.7%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Alfelali et al., 2020</td>
<td>0.09</td>
<td>0.0814</td>
<td></td>
<td>1.09</td>
<td>[0.93; 1.28]</td>
<td>28.8%</td>
<td>16.0%</td>
</tr>
<tr>
<td>MacIntyre et al., 2009</td>
<td>0.12</td>
<td>0.3335</td>
<td></td>
<td>1.13</td>
<td>[0.59; 2.18]</td>
<td>1.7%</td>
<td>3.9%</td>
</tr>
<tr>
<td>Simmerman, et al., 2011</td>
<td>0.15</td>
<td>0.2296</td>
<td></td>
<td>1.16</td>
<td>[0.74; 1.82]</td>
<td>3.6%</td>
<td>6.8%</td>
</tr>
<tr>
<td>MacIntyre et al., 2015</td>
<td>0.63</td>
<td>0.6386</td>
<td></td>
<td>1.88</td>
<td>[0.54; 6.57]</td>
<td>0.5%</td>
<td>1.2%</td>
</tr>
</tbody>
</table>

**Fixed effect model**

\[
\text{OR} = 0.88 \quad [0.80; 0.95] \quad 100.0\% \\
\text{Random effects model} \\
\text{OR} = 0.85 \quad [0.74; 0.98] \quad 100.0\%
\]

**Prediction interval**

\[
[0.50; 1.15]
\]