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

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

[1]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland

[2]Broad Institute of MIT and Harvard, Cambridge, MA, USA

[3]Stanford University School of Medicine, Palo Alto, CA, USA

[4]Helsinki Sleep Clinic, Vitalmed Research Center

[5]Department of Clinical Neurosciences, Clinicum, University of Helsinki, Helsinki, Finland

[1]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland

[2]Broad Institute of MIT and Harvard, Cambridge, MA, USA

[6]Helsinki University and Helsinki University Hospital, Clinic of Gastroenterology Helsinki, Finland

[7]Newcastle University Business School, Newcastle-upon-Tyne, United Kingdom

[8]Population Research Institute, Väestöliitto – The Family Federation of Finland

[9]The Wharton School of the University of Pennsylvania, Philadelphia, PA, USA

[10]Department of Medical Ethics and Health Policy, The Perelman School of Medicine, Philadelphia, PA, USA

*This version: December 23, 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

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, and Web of Science were searched for randomized controlled trials investigating the effect of face masks on respiratory infections published by November 18th, 2020. We followed PRISMA guidelines.

Eligibility criteria for selecting studies. Randomized controlled trials investigating face masks in respiratory infections across different exposure settings. Two reviewers performed the search, extracted data, and assessed the risk of bias. Random effects meta-analysis with risk ratio, adjusted odds ratios, and number needed to treat were performed. Findings by source control or wearer protection, age groups, exposure settings, and role of non-compliance were evaluated.

Results. Seventeen studies were included, (N=11,601 cases and N=10,286 controls, follow-up from 4 days to 19 months). Fourteen trials included adults and children and three trials included children only. Twelve studies showed non-compliance in treatment and eleven in control group. Four studies supported the use of face masks. Meta-analysis across all studies with risk ratios found no association with number of infections (RR=0.957 $[0.810 - 1.131]$, p=0.608). Meta-analysis using odds ratios adjusted for age, sex, and vaccination (when available) showed protective effect of face masks (OR=0.850 $[0.736 - 0.982]$, p=0.027). Subgroup meta-analysis with adjusted odds ratios found a decrease in respiratory infections among adults (14 studies, OR $= 0.829$ $[0.709 - 0.969]$, p=0.019) in source control setting (OR $= 0.845$ $[0.7375 - 0.969]$, p=0.0159) and when face masks were used together with hand hygiene OR $= 0.690$ $[0.568 - 0.838]$, p=0.0002). Overall between-study heterogeneity was large also in the subgroup analyses.

Conclusion. Despite the large between study heterogeneity, compliance bias and differences by environmental settings, the findings support the use of face masks to prevent respiratory infections.

PROSPERO registration number CRD42020205523.

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 [8–10]. 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 evidence on the efficacy of face masks use among the general public 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 arm [12–23] or in the control arm [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 randomized controlled trials, 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 examine the role of non-compliance in treatment and control arms, study whether the findings are affected by source control or wearer

protection, and the role of hand hygiene guidance on face masks.

2 Methods

This systematic review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses [31]. 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. The searches were limited 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.

We included two summary statistics measuring the risk of infection to our analysis. First is the unadjusted relative risk for infection which we computed using the raw numbers of infected in treatment arms which were provided in the studies. However, it is possible that omitted variables such as age, sex, vaccination or other factors such as correlation of outcomes within household affect the outcomes. Because almost none of the studies provided relative risk estimates adjusted for covariates, we used adjusted odds ratios in our secondary analyses if they were available in the studies. We discuss the limitations of these different summary statistics (odds ratio, risk ratio, for example) used in meta-analyses in detail in Section 4.3.

We excluded interventions that compared different types of face masks to each other (in which the comparison arm 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 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 (N per arm and odds ratio effect sizes with respective standard errors); (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. We contacted the corresponding author of [17] to confirm our interpretation of their data and also asked if they were able to provide the raw numbers for source control estimates. Unfortunately the data was not available anymore.

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 [32]. 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 arm and control arm. 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. 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].

2.5 Data analysis

The results for all the outcomes were expressed as risk ratios (RR), odds ratios (OR), and 95% confidence intervals for the effect estimates. 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.

We assessed heterogeneity and quantified statistical inconsistency by using the 2 test and the I2 statistic, respectively [33]. We used stratified meta-analyses to explore heterogeneity in the effect estimates according to: source control and wearer protection, 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, metaphor, and dmetar packages in R version 4.02 and meta package in Stata version 16.

Patient and public involvement

Our study research was not informed by patient and public involvement because we use secondary data. Patients or the public were not involved in the design, conduct, reporting or dissemination plans of our research.

3 Results

3.1 Search results

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).

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

Supplementary table 2 summarizes the characteristics of each study. Altogether, these studies included 11,601 study participants in the treatment arm and 10,286 in the control arm. The duration of follow-up varied from 4 days to 19 months. Nine of the seventeen studies reported raw values for wearer protection estimates exclusively [13–16]. Seven of the seventeen studies reported raw values for source control estimates exclusively [17–20, 22–25]. For study [19][19] raw values for source control were not available.

Studies 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].

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 arm 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. Some trials had a separate hand hygiene, face mask or hand hygiene with face mask arm [13, 14, 18, 19] in which the treatment arm the intervention included also a hand sanitizer or instruction to wash hands.

Three studies [13, 14, 24] found a protective effect of face masks in the intention-to-treat analysis and one in the analysis of secondary transmission rate [19]. Two of these studies had a follow-up length of 6 weeks [13, 14] and third one up to 19 months [19]. These three studies [13, 14, 19] had also 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 control 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 so that transmission and infection has not already happened before adopting mask use.

3.4 Risk of bias across the studies

Table 1 summarizes the risk of bias on the study level. The observed bias was low or unclear in the majority of the 17. As discussed in our methods section, we did not exclude studies based on bias. 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 arm 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 fully blinded in any of the 17 studies (Table 1).

3.5 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 2). 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: $\beta = 0.36$, se $= 0.52$, p $= 0.496$.

3.6 Face masks and respiratory infections

There were in total 1,542 events among the treatment arms (N total treatment = 11, 601) and 1, 345 events in control arms (N total control = 10, 286). The meta-analysis using unadjusted risk ratios (from raw N values) across all 17 studies finds no association between face mask intervention and respiratory infections (RR = 0.957 [0.810 − 1.131], p = 0.6086) (Figure 3a). The analysis using covariate adjusted odds ratios shows a protective effect of face masks (OR = 0.850 [0.736 − 0.982] p = 0.027, Figure 3b).

The overall effect of face masks consists of the combination of source control and personal protection for the mask wearer (wearer protection). We investigated this mechanism by performing a subgroup analysis by categorizing the studies exclusively based on their estimated effects on source control or wearer protection. Subgroup meta-analysis based on the unadjusted risk ratios was inconclusive (wearer protection RR = 0.868 [0.720 − 1.047], p = 0.139; source control RR = 1.0054 [0.685 − 1.477], p = 0.978, Figure 3c). Secondary analysis using adjusted odds ratios was also inconclusive on wearer protection (OR = 0.836 [0.661 − 1.056] p = 0.133) but indicates a protective effect as a source control (OR = 0.845 [0.738 − 0.969] Figure 3d). Similarly, raw values did not associate with fewer infections in individuals over 15 years of age (RR = 0.920 [0.774 − 1.094], p = 0.343, Figure 4) whereas adjusted ORs did (RR = 0.829 [0.709 − 0.969] p = 0.019).

In addition to the substantial differences between the analysis using unadjusted and adjusted effect estimates, it is also possible that individual studies bias the estimates of the meta-analysis. Thus, we performed a meta-analysis through a leave-one-out analysis to examine if one study causes a systematic association to a given direction. The effect sizes were all systematically at RR < 1. However, one of the largest studies, [15], had a significant level of non-compliance in the treatment arm with 49% of the controls using face masks; and the second largest study which had the longest follow-up time, [19], we did not have access to secondary transmission values.

We investigated the importance of non-compliance on the results by performing a subgroup meta-analysis without the studies with non-compliance of over 10% in the control arm. This showed a protective effect of face masks in adjusted ORs but not with raw estimates (RR = 0.877 [0.686 − 1.121], OR = 0.7641 [0.682 − 0.856] p < 0.0001) indicating that non-compliance has weakened the power to observe an association in these trials.

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 4). In the random effects meta-analysis in the hospital or household settings association was seen with adjusted ORs but not with raw effect estimates (RR = 1.207 [1.114 − 1.307], p = 0.219, OR = 0.861 [0.752 − 0.987], p = 0.0318). The effect was similar in the community settings although statistically insignificant (RR = 0.838 [0.689 − 1.012], p = 0.077, Figure 4). It is possible that these large confidence intervals in the community setting result from non-compliance (between 10% and 50% of the control groups using masks) in three out of the six studies that assessed community transmission and from relatively low compliance in the treatment groups. Indeed, meta-regression adjusting for face mask use in the control arm showed statistically significant association at the population setting (RR = 0.86 [0.79- 0.94], p = 0.0007).

In the twelve of the trials including adults, the intervention consisted solely of face mask use while in six the intervention included also guidance on appropriate hand hygiene together with the face mask use. This subgroup analysis for the face mask with hand hygiene guidance showed effect with adjusted ORs but not with raw estimates (RR = 1.020 [0.738 − 1.408], p = 0.900, OR =0.690 [0.568 − 0.838], p = 0.0002, Figure 5).

3.7 Number needed to treat

We approximate the effect of masks on population health by exploring the number needed to treat (NNT), 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.

4 Discussion

4.1 Main findings

This systematic review and meta-analysis of 17 randomized controlled trials examined whether face mask intervention can prevent respiratory infections. 4 out of 17 studies supported the use of

masks. In addition, the analysis of the adjusted odd ratios showed a protective effect of face masks across different settings. However, the pooled effect using the unadjusted effects based on raw numbers versus adjusted estimates were inconsistent. This together with remaining heterogeneity even in subgroup analyses indicate large variation across these studies that needs to be taken into account both in the interpretations of the results and when designing future research on face masks. The protective effect of face masks was observed both with face mask use alone and when face masks were combined with appropriate hand hygiene. This result is aligned with the current evidence that face masks are most efficient when used together with appropriate hand hygiene and other protective measures. 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 randomized controlled trials, case-control studies and cohort studies [30] or cloth masks [29] in non-healthcare settings. The findings from our systematic review and meta-analysis are 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 face mask treatment arm or the use of face masks in the control arm, which may bias the estimates towards the null hypothesis. This is also confirmed by our sensitivity analysis.

In addition, as the effects with hand hygiene seems to be stronger than with mask use alone, one 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 [7]. Also in our study, the effect of masks was further accentuated when combined with complementary measures, such as improved hand hygiene [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 [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 [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 populations 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 (influenza-like illness, 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 discussed, we also did not have access to the raw values for secondary transmission (source control) in one of the studies [19]. 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

This systematic review and 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.

Contributorship statement

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.

Funding

Ollila and Laine 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.

Data sharing

We provide raw numbers and adjusted odds ratios as supplementary information along with the code that was used to compute the associations.

References

[1] Cutler DM and Summers LH. The COVID-19 Pandemic and the $16 Trillion Virus. *JAMA*, 324(15):1495–1496, 2020. doi: 10.1001/jama.2020.19759.

[2] Murray EJ. Epidemiology's Time of Need: COVID-19 Calls for Epidemic-Related Economics. *J Econ Perspectives*, 34(4):105–20, 2020. doi: 10.1257/jep.34.4.105.

[3] Liu J, Liao X, and Qian S et al. Community Transmission of Severe Acute Respiratory Syndrome Coronavirus 2, Shenzhen, China, 2020. *Emerg Infect Dis*, 26(6):1320–1323, 2020. doi: 10.3201/eid2606.200239.

[4] Sharma SK, Mudgal SK, Panda P, Gupta P, and Aggarwal P. COVID–19: Guidance Outlines on Infection Prevention and Control for Health Care Workers. *Indian J Community Health*, 32(1):8–14, 2020. doi: 10.47203/IJCH.2020.v32i01.004.

[5] World Health Organization. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). `https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf`, 2020.

[6] van Doremalen N, Bushmaker T, and Morris DH et al. Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1. *N Engl J Med*, 382(16):1564–1567, 2020. doi: 10.1056/NEJMc2004973.

[7] Chu D, Akl E, and Duda S et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet*, 395(10242):1973–1987, 2020. doi: 10.1016/S0140-6736(20)31142-9.

[8] Kai D, Goldstein G-P, Morgunov A, Nangalia V, and Rotkirch A. Universal Masking is Urgent in the COVID-19 Pandemic: SEIR and Agent Based Models, Empirical Validation, Policy Recommendations, 2020.

[9] Stutt ROJH, Retkute R, Bradley M, Cilligan CA, and Colvin J. A modelling framework to assess the likely effectiveness of facemasks in combination with 'lock-down' in managing the COVID-19 pandemic. *Proc R Soc A*, 476(2238):20200376, 2020. doi: 10.1098/rspa.2020.0376.

[10] Hoertel N, Blachier M, and Blanco C et al. A stochastic agent-based model of the SARS-CoV-2 epidemic in France. *Nat Med*, 2020. doi: 10.1038/s41591-020-1001-6.

[11] Miyazawa D and Kaneko G. Face mask wearing rate predicts country's COVID-19 death rates: with supplementary state-by-state data in the United States. *medRxiv*, 2020. doi: 10.1101/2020.06.22.20137745.

[12] Abdin EZ, Choudhry AJ, and Al-Naji A. Effect of use of Face mask on Hajj related Acute Respiratory Infection among Hajjis from Riyadh - A Health Promotion Intervention Study. *Saudi Epidemiology Bulletin*, 12(4):27–28, 2005.

[13] Aiello AE, Murray GF, and Perez V et al. Mask Use, Hand Hygiene, and Seasonal Influenza-Like Illness among Young Adults: A Randomized Intervention Trial. *J Infect Dis*, 201(4): 491–498, 2010. doi: 10.1086/650396.

[14] Aiello AE, Perez V, Coulborn RM, Davis BM, Uddin M, and Monto AS. Facemasks, Hand Hygiene, and Influenza among Young Adults: A Randomized Intervention Trial. *PLoS ONE*, 7(1):1–8, 2012. doi: 10.1371/journal.pone.0029744.

[15] Alfelali M, Haworth E., and Barasheed O et al. Facemask against viral respiratory infections among Hajj pilgrims: A challenging cluster-randomized trial. *PLoS ONE*, 15(10):1–20, 2020. doi: 10.1371/journal.pone.0240287.

[16] Bundgaard H, Bundgaard JS, and Raaschou-Pedersen DET et al. Effectiveness of Adding a Mask Recommendation to Other Public Health Measures to Prevent SARS-CoV-2 Infection in Danish Mask Wearers. *Ann Intern Med*, Forthcoming, 2020. doi: 10.7326/M20-6817.

[17] Cowling BJ, Fung ROP, and Cheng CKY al et. Preliminary Findings of a Randomized Trial of Non-Pharmaceutical Interventions to Prevent Influenza Transmission in Households. *PLoS ONE*, 3(5), 2008. doi: 10.1371/journal.pone.0002101.

[18] Cowling BJ, Chan K-H, and Fang VJ et al. Facemasks and Hand Hygiene to Prevent Influenza Transmission in Households. *Ann Intern Med*, 151(7):437–446, 2009. doi: 10.7326/0003-4819-151-7-200910060-00142.

[19] Larson EL, Ferng Y-H, Wong-McLoughlin J, Wang S, Haber M, and Morse SS. Impact of Non-Pharmaceutical Interventions on URIs and Influenza in Crowded, Urban Households. *Public Health Rep*, 125(2):178–191, 2010. doi: 10.1177/003335491012500206.

[20] MacIntyre CR, Cauchemez S, and Dwyer DE et al. Face Mask Use and Control of Respiratory Virus Transmission in Households. *Emerg Infect Dis*, 15(2):233–241, 2009. doi: 10.3201/eid1502.081167.

[21] MacIntyre CR, Zhang Y, and Chughtai AA et al. Cluster randomised controlled trial to examine medical mask use as source control for people with respiratory illness. *BMJ Open*, 6: 1–9, 2016. doi: 10.1136/bmjopen-2016-012330.

[22] Suess T, Remschmidt C, and Schink SB et al. 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):1–16, 2012. doi: 10.1186/1471-2334-12-26.

[23] Simmerman JM, Suntarattiwong P, and Levy J et al. Findings from a household randomized controlled trial of hand washing and face masks to reduce influenza transmission in Bangkok, Thailand. *Influenza and Other Respiratory Viruses*, 5(4):256–267, 2011. doi: 10.1111/j.1750-2659.2011.00205.x.

[24] Barasheed O, Almasri N, and Badahdah A-M et al. Pilot Randomised Controlled Trial to Test Effectiveness of Facemasks in Preventing Influenza-like Illness Transmission among Australian Hajj Pilgrims in 2011. *Infectious Disorders - Drug Targets*, 14:110–116, 2014. doi: 10.2174/1871526514666141021112855.

[25] Canini L, Andréoletti L, and Ferrari P et al. Surgical Mask to Prevent Influenza Transmission in Households: A Cluster Randomized Trial. *PLoS ONE*, 5(11):1–6, 2010. doi: 10.1371/journal.pone.0013998.

[26] Jacobs JL, Ohde S, Takahashi O, Tokuda Y, Omata F, and Fukui T. 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, 2009. doi: 10.1016/j.ajic.2008.11.002.

[27] MacIntyre CR, Wang Q, and 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 and Other Respiratory Viruses,*, 5(3):170–179, 2011. doi: 10.1111/j.1750-2659.2011.00198.x.

[28] MacIntyre CR, Seale H, and Dung TC et al. A cluster randomised trial of cloth masks compared with medical masks in healthcare workers. *BMJ Open*, 5:1–10, 2015. doi: 10.1136/bmjopen-2014-006577.

[29] Sharma SK, Mishra M, and Mudgal SK. Efficacy of cloth face mask in prevention of novel coronavirus infection transmission: A systematic review and meta-analysis. *J Educ Health Promot*, 9(1):192–200, 2020. doi: 10.4103/jehp.jehp_533_20.

[30] Liang M, Gao L, and Cheng C et al. Efficacy of face mask in preventing respiratory virus transmission: A systematic review and meta-analysis. *Travel Med Infect Dis*, 36, 2020. doi: 10.1016/j.tmaid.2020.101751.

[31] Moher D, Liberati A, Tetzlaff J, and Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*, 339, 2009. doi: 10.1136/bmj.b2535.

[32] Higgins JP, Altman DG, and Gøtzsche PC et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, 343:1–9, 2011. doi: 10.1136/bmj.d5928.

[33] Higgins JPT, Thompson SG, Deeks JJ, and Altman DG. Measuring inconsistency in meta-analyses. *BMJ*, 327(7414):557–560, 2003. doi: 10.1136/bmj.327.7414.557.

[34] Suess T, Remschmidt C, and Schink SB et al. Facemasks and intensified hand hygiene in a German household trial during the 2009/2010 influenza A(H1N1) pandemic: adherence and tolerability in children and adults. *Epidemiol Infect*, 139(12):1895–1901, 2011. doi: 10.1017/S0950268810003006.

[35] Jefferson T, Foxlee R, and Del Mar C et al. Physical interventions to interrupt or reduce the spread of respiratory viruses: systematic review. *BMJ*, 336:77–80, 2020. doi: 10.1136/bmj.39393.510347.BE.

[36] Taylor F, Huffman MD, and Macedo AF et al. Statins for the primary prevention of cardiovascular disease. *Cochrane Database Syst Rev.*, 2013. doi: 10.1002/14651858.CD004816.pub5.

Appendix

Study criteria

Table A1: Study criteria

1. Population: General population and health care personnel in risk of contracting infectious diseases.
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).
12. Excluded: Animal studies.

A. Search terms

Our literature search can be replicated using the following protocol.

B. 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

C. 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] OR "Central Nervous System Viral Diseases"[Mesh] OR "Influenza,

Human”[Mesh] OR ”Influenza A Virus, H5N1 Subtype”[Mesh] OR ”Influenza A Virus, H1N1 Subtype”[Mesh] OR ”Influenza A virus”[Mesh] OR ”SARS Virus”[Mesh] OR ”viral infection” OR ”corona virus” or ”swine flu” or ”MERS”) OR ”COVID-19” [Supplementary Concept]) AND (”Systematic review” OR ”cohort study” OR ”case-control” OR ”Randomized Controlled Trial” [Publication Type] OR ”Randomized Controlled Trials as Topic”[Mesh] OR ”Controlled Clinical Trial” [Publication Type] OR ”Meta-Analysis” [Publication Type] OR ”Meta-Analysis as Topic”[Mesh] OR ”Network Meta- Analysis”[Mesh])

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.

D. 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.

			Blinding				Non-compliance	
	Sequence generation	Allocation concealment	Participants and personnell	Outcome assesment	Incomplete outcome data	Selective reporting	Treatment group	Control group
Abdin et al., 2005	Unclear risk	Unclear risk	Unclear risk	Unclear risk	Low risk	Low risk	Unclear risk	Unclear risk
Aiello et al., 2010	Unclear risk	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Unclear risk	Low risk
Aiello et al., 2012	Low risk	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Unclear risk	Low risk
Alfelali et al., 2020	Unclear risk	Unclear risk	Unclear risk	Unclear risk	Low risk	Low risk	Unclear risk	High risk
Barasheed et al., 2014	Low risk	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Low risk	Unclear risk
Bungaard et al. 2020	Low risk	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Unclear risk	Low risk
Canini et al., 2010	Low risk	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	Unclear risk
Cowling et al., 2008	Low risk	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Unclear risk	Unclear risk
Cowling et al., 2009	Low risk	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Unclear risk	Unclear risk
Jacobs et al., 2009	Low risk	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Low risk	Unclear risk
Larson et al., 2010	Low risk	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Unclear risk	Low risk
Macintyre et al., 2009	Low risk	Low risk	Low risk	Unclear risk	Low risk	Low risk	Unclear risk	Low risk
Macintyre et al., 2011	High risk	Unclear risk	Unclear risk	Unclear risk	Low risk	Low risk	Low risk	Unclear risk
Macintyre et al., 2015	Low risk	Unclear risk	Unclear risk	Unclear risk	Low risk	Low risk	Low risk	Unclear risk

MacIntyre et al., 2016	Low risk	Low risk	Unclear risk	Unclear risk	Low risk	Low risk	Unclear risk	Unclear risk
Simmerman et al., 2011	Low risk	Low risk	Low risk	Unclear risk	Low risk	Low risk	High risk	Unclear risk
Suess et al., 2012	Low risk	Low risk	Low risk	Unclear risk	Low risk	Low risk	High risk	Low risk

Table 1. 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 captions

Figure 1. Prisma flow diagram of included articles

Figure 2. Review authors' judgement of each of risk-of-bias item as percentage across the included studies.

Figure 3. Pooled effect across all 17 studies using a) raw estimates and relative risk b) adjusted odds ratios when available c) raw estimates examining source control and wearer protection separately and d) adjusted odds ratios examining source control and wearer protection separately. TE = treatment effect and seTE = standard error of treatment effect. Fixed and Random effects pooled estimates are shown.

Figure 4. Subgroup analysis in a) adults and in children using raw estimates and relative risk b) adults and in children using adjusted odds ratios when available c) community setting and indoor environment separately using raw estimates and d) adjusted odds ratios examining community setting vs indoor environment using raw estimates separately. TE = treatment effect and seTE = standard error of treatment effect. Fixed and Random effects pooled estimates are shown.

Figure 5. Subgroup analysis in a) face mask together with instructed hand hygiene using raw estimates b)) face mask together with instructed hand hygiene using adjusted odds ratios when available. TE = treatment effect and seTE = standard error of treatment effect. Fixed and Random effects pooled estimates are shown.

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.

Supplementary Table. Study characteristics

R code for running analyses.

Values from original articles to use alongside the R code.

Figure 1

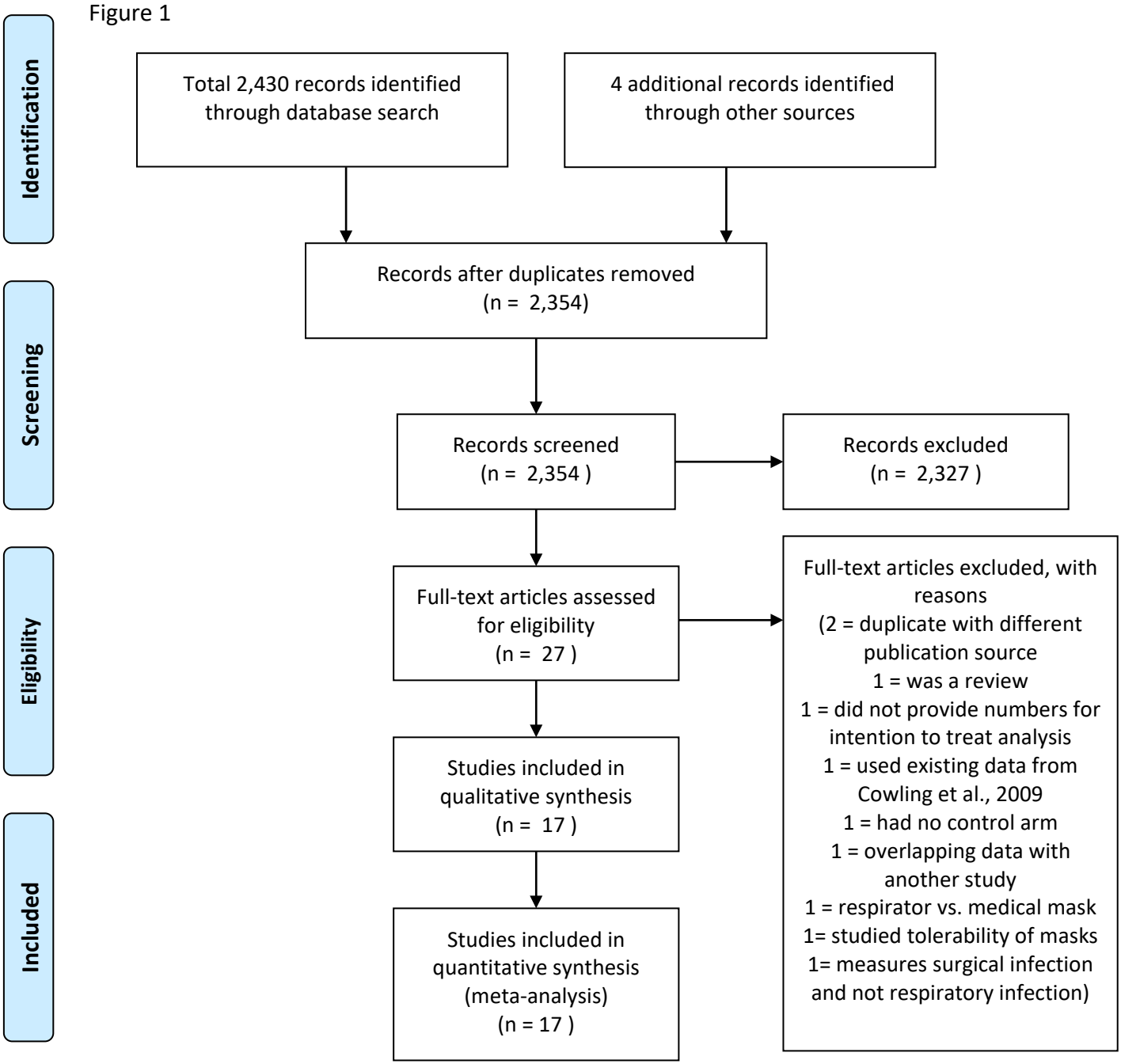

Identification
Screening
Eligibility
Included
Total 2,430 records identified through database search
4 additional records identified through other sources
Records after duplicates removed (n = 2,354)
Records screened (n = 2,354)
Records excluded (n = 2,327)
Full-text articles assessed for eligibility (n = 27)
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)
Studies included in qualitative synthesis (n = 17)
Studies included in quantitative synthesis (meta-analysis) (n = 17)

Figure 2

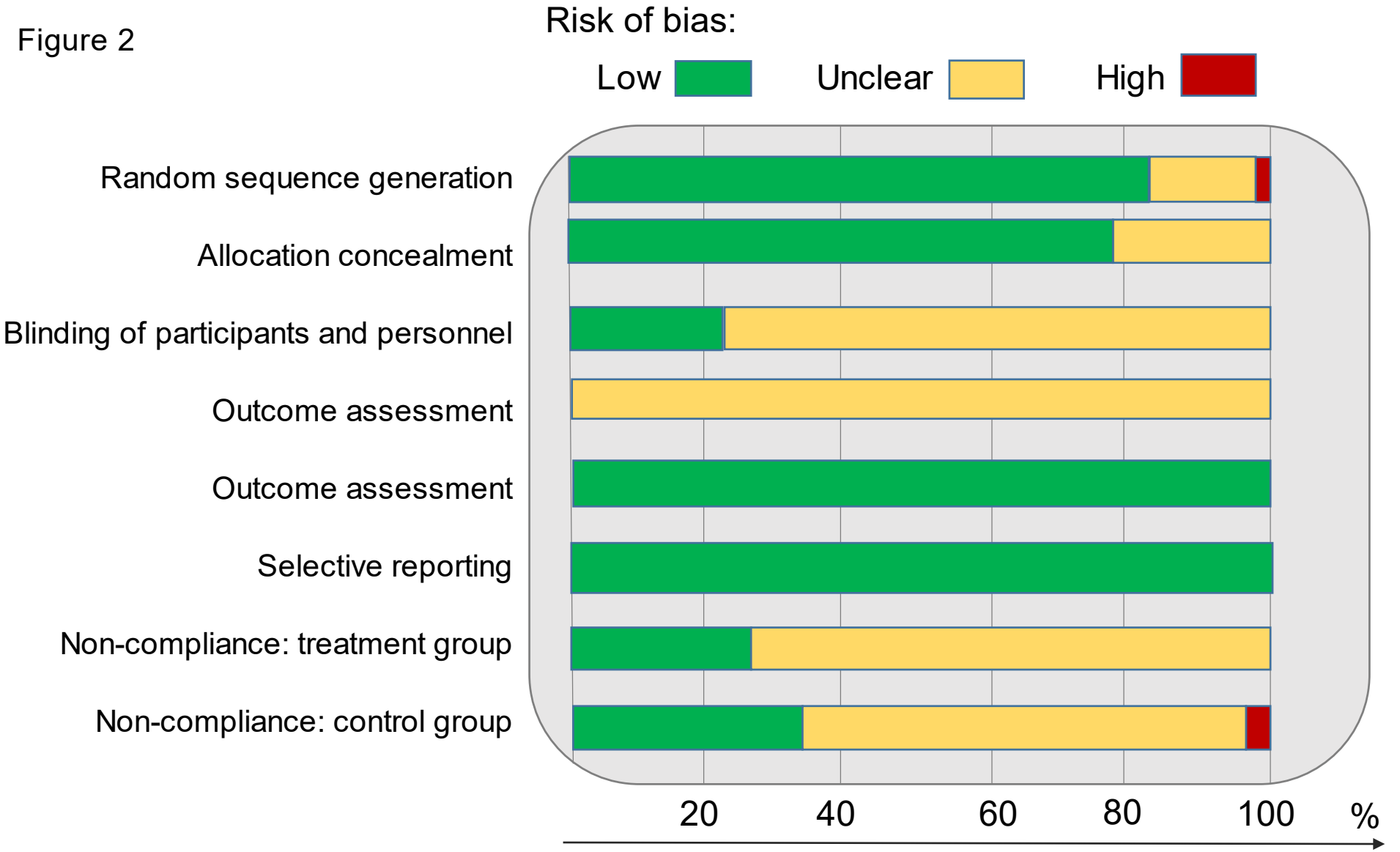

Risk of bias:
Low
Unclear
High
Random sequence generation
Allocation concealment
Blinding of participants and personnel
Outcome assessment
Outcome assessment
Selective reporting
Non-compliance: treatment group
Non-compliance: control group
20
40
60
80
100
%

Table 2. Characteristics of included studies.

Article	Year	Setting	Region, Country	N	Age	Randomization level	Follow-up	Type of mask and other interventions	Disease and/or index case definition	Compliance	Main findings
Abdin et al., 2005	2004	Pilgrims	Hajj, Saudi Arabia	**Total N = 995** - 446 no intervention - 292 health education - 257 health education and facemask - Adults	Mean age of 35.3 years (SD ±11.72) 425 (43%) females with mean age 34.7 years (SD ± 13.71).	Tents	9 days + one week followup for symptoms	**- Control group** **- Control with health education: healh education** leaflet **- Mask group: E**ducation leaflet and facemask	**At least one of the constitutional symptoms** fever, headache, myalgia **Along with one of the local symptoms** runny nose, sneezing, throat pain, cough with or without Sputum, difficulty in breathing Symptoms by self-report questionnaire.	**Wearing the mask sometimes and wearing the mask always were considered as compliant** 33.6% of controls used facemasks 51.7% of health education group used facemasks 81.3% of facemask group used facemask	- Facemask use associated with treatment allocation group but not with rate of infection.
Aiello et al., 2010	2006 - 2007	University students	Michigan, USA	**Total N = 1,437** - Control = 552 - Facemask = 378 - Facemask and Hand hygiene = 367 - Adults	Adults over 18 years Mean age 18.7 years SD =0.8	University residence halls	6 weeks	- **All participants:** Basic hand hygiene education and cough etiquette through an email video link and the study Web site. - **Facemask and hand hygiene group participants:** Written materials detailing appropriate hand sanitizer and mask use - Mask groups: Medical procedure masks with ear loops (TECNOL procedure masks; Kimberly-Clark) - **Mask and hand hygiene group:** Resealable plastic bags for mask storage when not in use and for disposal - **All participants** received alcohol-based hand sanitizer (portable 2 oz squeeze bottle; 8 oz pump)	**ILI was defined as presence of cough and at least 1 constitutional symptom** - fever/feverishness, chills, or body aches] - Phone contact with the nursing staff to assess for ILI symptoms. - During scheduled participant visits, study nurses ascertained date of illness onset, temperature, use of antipyretics, and reported symptoms (cough, feverishness, chills, body aches, headache, nasal congestion, and sore throat).	**Mask compliance:** Average number of mask hours per day during the past week. Hand hygiene complianc: Assessed by reported use of hand sanitizer in the face mask/hand hygiene intervention group. Average number of times the hand sanitizer was used per day in the past week and the amount used. Participants were observed for instances of hand sanitizer use. Mask only group wore their mask 3.92 hours per day versus 2.99 hours per day in the mask and hand hygiene group. Mask only group washed their hands 8.18 times per dayversus 6.11 times per day in the mask and hand hygiene group. The control group washed their hands on average 8.75 times per day. Mask and hand hygiene group washed their hands significantly fewer times per day than the control group from weeks 2 through 4 only. On average, the mask only group washed their hands for 23.15 seconds per day versus 20.65 seconds per day in the mask and hand hygiene group.	- Reductions in ILI during weeks 4–6 in the mask and hand hygiene group. - Face mask use alone associated with lower ILI in raw model but not with adjusted estimates - Cumulative analyses were not significant
Aiello et al., 2012	2007 - 2008	University students	Michigan, USA	**Total N = 1,178** - Control = 396 - Facemask = 420 - Facemask and hand hygiene = 362 - Adults	Adults over 18 years Mean age 18.95 years SD= 0.9	University residence halls	6 weeks	- **Mask and mask with hand hygiene group:** Weekly packets of mask supplies: TECNOLTM procedure masks, Kimberly-Clark, Roswell GA - Plastic bags for storage and daily disposal. - **Facemask and hand hygiene intervention:** 2 oz squeeze bottle, 8 oz pump bottle with 62% ethyl alcohol in a gel base - The control group did not receive an intervention.	**Symotoms of ILI** - Presence of cough and at least one or more of - Fever/feverishness, chills, or body aches - Contact information of clinical research staff for illness assessment - Clinical research staff recorded the date of illness onset, body temperature, use of anti-pyretics, and reported symptoms. **Laboratory-confirmed influenza in a subset** - Throat swab specimens were tested for influenza A or B using real-time polymerase chain reaction (Rt-PCR).	**Mask use compliance** Controls did not use facemasks **Subjects** - Face mask and hand hygiene group wore their mask, on average, 5.08 hours per day, mask only group (5.04 hours per day). - No significant difference in mask use between the two interventions was observed throughout the study. - Face mask and hand hygiene group reported an average use of hand sanitizer 4.49 times per day. - Mask only group reported an average use of hand sanitizer 1.29 times per day. - Control group reported use of 1.51 times per day. - Face mask and hand hygiene group used hand sanitizer significantly more often compared to subjects in either the mask only or control groups.	- Reduction in the rate of ILI was observed in weeks 3 through 6 - Cumulative results were not statistically significant
Alfelali et al., 2020	2013, 2014, 2015	Pilgrims	Hajj, Saudi Arabia	**Total N = 381 tents and 7,687 participants** - Control 169 tents, 3823 participants - Facemask: 149 tents, 3864 participants - Adults	Adults over 18 years 18 - 95 years median 34 mean 37and SD = 12 years	Tents	4 days	- **Mask group** was provided with 50 surgical facemasks: 3M™ Standard Tie-On surgical mask, Cat No: 1816 - Verbal and printed instructions, demonstration of appropriate facemask usage - No mask in controls	**Symptoms:** -respiratory symptoms were recorded daily. Fever plus one respiratory symptom, or two or more respiratory symptoms without fever. Pilgrims from participating countries staying in their respective tents. Age over 18 years. The participant should be able to provide signed informed consent. Excluded: Age < 18 years. Participation in another clinical trial investigating a medical intervention that may interfere with the study outcome measures, like laboratory-confirmed viral RTI. Any known contraindication to mask use (e.g., allergy to standard [surgical] mask materials). Refusal/or inability to sign the consent form.	**Mask use compliance** 49.2% used facemask in the control arm 72.4% used facemask in the test arm	- No conclusive evidence on facemask efficacy most likely due to poor adherence to protocol
Barasheed et al., 2014	2011	Pilgrims	Hajj, Saudi Arabia	**Total N = 164 in 22 tents** - Controls; 10 tents; 89 in the control group - Facemask; 12 tents; 75 individuals - Adults	Age over ≥ 15 Age median in control arm 41.6 (range 17-72) Age median in mask arm 48 (range 19-80)	Tents	7 days, 5 days recorded symptoms	- **Facemask group** provided - Plain surgical facemask: 3MTM Standard Tie-On Surgical Mask, Cat No: 1816 - No mask in controls	**ILI was defined as subjective (or proven) fever plus one respiratory symptom** Dry or productive cough, runny nose, sore throat, shortness of breath - cases: Those reporting new onset of ILI - contacts: pilgrims who shared the same tent and slept in an adjacent bed within 2 meters to cases	**Mask use compliance** 76% in the mask group 12% in the control group used masks.	- Less contacts became symptomatic in the mask tents
Bundgaard et al., 2020	2020	Population	Denmark	**Total N = 6024** - 3030 individuals in mask group - 2994 in control group - Mask group was recommended to wear a facemask	Individuals over 18 years of age	Population	1 month	- Mask group was provided 50 three-layer, disposable, surgical face masks with ear loops (TYPE II EN 14683 [Abena]; filtration rate, 98%; made in China) - Aged 18 years or older - No current or prior symptoms or diagnosis of COVID-19 - Reported being outside the home among others for at least 3 hours per day and who did not wear masks during their daily work	**COVID-19 infection** If symptoms of illness occurred the study participants collected a oropharyngeal/nasal swab. Positive laboratory test for COVID-19 was considered an infection.	**Among face mask group** - 7% non adherent - 46% of participants wore the mask as recommended, 47% predominantly as recommended **Control group** - The exact values are not reported/unknown - No mask mandate in Denmark during the study suggest low non compliance in controls	- No significant association with face mask use.
Canini et al., 2010	2008-2009	Household	Ile de France, Aquitaine and Franche-Comte´, France	**Total N = 105 households and 306 contacts** - Controls N = 158 - Facemask N = 148 - Adults and children	Median age in mask arm 25 SD 16 years Median age in control arm 28 SD 16 years	Household stratified according to age of the index patient: under or over 15 years	7 days, 5 days active intervention	- Only **index case had facemask**. - Surgery masks with earloops, 3 plys, anti fog. - No mask in controls	**Symptoms lasting less than 48 hours, combining temperature over 37.8C and cough, and a positive rapid test for influenza A.** Index case needed to be aged over 5 years old. The index patient had to be a priori the first and unique illness case in the household and be affiliated to the French national health insurance. Exclusion criteria: Index case had asthma or chronic obstructive pulmonary disease or hospitalized. Primary outcome case: household contacts who developed an ILI during the 7 days following inclusion. A temperature over 37.8C with cough or sore throat was used as primary clinical case-definition.	**Compliance** Reported low compliance The index patients reported wearing a total of 11 masks during 4 days with an average use of 2 masks per day and a duration of use of 3.7 hours a day.	- Intention to treat analysis did not show statistically significant benefit

Cowling et al, 2008	2007	Household	Hong Kong	**Total N = 259 households and 794 household members** - Control 74 households, 213 contacts - Facemask 22 households, 61 contacts - Hand hygiene 32 households, 92 - Adults and children	Index patient age range from 2 years to over 50 years Household contact from 0 to over 50 years	Household	9 days	- **Mask group:** Surgical mask - No mask in controls	**Fever 38'C, cough, sore throat, coryza, headache, malaise, chills, fatigue** - Residents aged at least 2 years - Reporting at least two symptoms of ILI - Living in a household with at least two other individuals - No reported ILI symptoms in the preceding 14 days	**Compliance** Over 25% household contacts in the face mask group did not wear a surgical mask at all Adherence to the face mask intervention was higher in the index subjects **Controls** Over 25% of the index cases in the control and hand hygiene intervention arms reported wearing masks at home of their own accord	- No significant differences between intervention arms - The secondary attack ratios were lower than anticipated
Cowling et al. 2009	2008	Household	Hong Kong	**Total N = 407 households and 1015 contacts** - Control 91 households, 346 contacts - Hand hygiene 85 households, 329 contacts - Facemask and hand hygiene = 93 households, 340 contacts - Adults and children	Median age 10 in control arm Median age 10 in mask arm	Household	6 days	**- Controls:** education on healthy diet and lifestyle **- Hand hygiene group:** instructions, liquid handsoap and hand sanitizer **- Mask and hand hygiene:** education and demonstratio on mask use, 50 surgical facemasks Techol the lite one (Kimberly-Clark), 75 pediatric masks	**Persons who reported at least 2 symptoms of acute respiratory illness** - Temperature higher or equal to 37.8C, cough, headache, sore thorat or myalgia - Had symptom onset within 49 hours - Lived in a household with at least two other people - None of whom had reported acute respiratory illness in the preceding 14 days.	**Compliance** Intervention groups reported higher adherence to the interventions than the control group. 50% of index patients in the facemask plus handhygiene group reported regular use of a surgical mask during follow-up Facemask adherence among household contacts was lower Adherence to the hand hygiene intervention seemed low compared with rates recommended in health care settings but was similar to rates in previous community settings **Controls** Some contamination between groups was observed, because both interventions were practiced to some degree in the control group	- No associaton in the total analysis - Significant reduction in RT-PCR confirmed infections where the intervention was applied within 36 hours of sympton onset - Hand hygiene and facemasks can reduce influenza virus transmission if impelemted early after symptom onset in an index patients. - Due to problems with adherence, effects in the study may tend toward a lower bound on the effects. - Authors highlight ways to improve adherence for future studies.
Jacobs et al., 2010	2008	Healthcare workers	Tokyo, Japan	**Total N = 33**, one dropped out - 15 in control group - 17 in face mask group - Adults	Median age in mask arm 35 SD = 14 Median age in control arm 36 SD = 9.6	Hospitals by job category: nurses, doctors, and comedical personnel	77 days	**Study arms** - Facemask group had masks - Surgical mask MA-3 (Ozu Sangyo, Tokyo, Japan) - No mask in controls	**Mofidied scale to determine illness** - 8 symptoms of infection on a 4-point scale (0, none; 1, mild; 2, moderate; 3, severe; for fever: 0, absent; 1, present). - Criterion for infection was a 2-day total symptom score greater than 14 (modified Jackson criteria). - Exclusion criteria: self identification of conditions predisposing to infections or taking antibiotics	**Compliance** Compliance was 84.3% in the treatment group. 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 (eg, surgical nurse in the operating room). Subjects' face mask wearing was controlled only in the hospital. Behavior outside of the hospital was not measured nor was frequency of replacing face masks.	- Face mask use in health care workers did not provide benefit in terms of cold symptoms
Larson et al., 2010	2006 - 2008	Household	New York, USA	**Total N = 509 households, 2788 participants** - Control (education); 211 households N = 904 (dropped 26 households) - Hand sanitizer; 205 households N = 946 (dropped 21 households) - Facemask and hand hygiene; 201 households, N = 938 (dropped 19 households) - Adults and children	The majority of index cases were adults over 18 years of age. Age range of household contacts range from 0 to over 65 years.	Household	19 months	**- Hand sanitizer group:** educational materials and hand sanitizer: Purell®, Johnson & Johnson, Morris Plains, New Jersey in large (8- and 4-ounce) and small (1-ounce) containers **-Facemask and hand hygiene group:** Procedure Face Masks for adults and children, Kimberly- Clark, Roswell, Georgia. Instructions for use.	**Infection based on symptoms** - Six symptoms reported at least twice per week: rhinorrhea (runny nose), sore throat, cough, muscle aches, fever, and headache. When an ILI was reported, an alert was electronically sent to the project staff, who immediately contacted the reporting household. A member of the research team was then deployed to make a home visit within 24 to 48 hours to obtain a sample for laboratory testing for influenza. - Secondary case onset required within five days following the index case	**Compliance** - 50% of individuals in the mask group reported using mask within 48h of disease onset. - Those who used masks at all reported a mean of two masks per day or ILI episode - Range: in cases from no mask use at all to nine masks per episode	- Statistically significant association with facemask preventing respiratory infections
MacIntyre et al., 2009	2006, 2007	Household	Sydney, Australia	**Total N = 143 households and 286 exposed adults** - Control; 50 families and 100 adult contacts (2 excluded) - Surgical mask; 47 families and 94 adult contacts - P2 mask; 46 families and 92 contacts - Children as index patients	Index case is a child between 0-15 years Contact over 15 years.	Household	5 days	**Study arms** - Surgical masks - Non–fit tested P2 masks - No masks in controls	**Index child had fever (temperature >37.8oC) and either a cough or sore throat** The household contained >2 adults >16 years of age and 1 child 0–15 years of age The child was the first and only person to become ill in the family in the previous 2 weeks - Index child was not admitted to the hospital **ILI in the household contact** **-** The presence of fever (temperature >37.8°C), feeling feverish or a history of fever, - >2 symptoms (sore throat, cough, sneezing, runny nose, nasal congestion, headache), or 1 of the symptoms listed plus laboratory confirmation of respiratory viral infection.	**Compliance** - 38% of the surgical mask users used the mask most or all of the time - 46% of the P2 mask group using most or all of the time - Other participants were wearing face masks rarely or never Adherence dropped to 31% and 25% by day 5 of mask use	- The intention-to-treat analysis showed no difference between arms. - <50% of participants were adherent with mask use - Adherent mask users had a significant reduction in the risk for clinical infection
MacIntyre et al., 2011	2008 - 2009	Healthcare workers	Beijing, China	**Total N = 1441 Healthcare workers in 15 hospitals** - control group = 9 hospitals, convenience control - medical mask, 5 hospitals N=492 - N95 non fit tested, 5 hospitals N=488 - N95 fit tested, 5 hospitals N=461 - Adults	Individuals over 18 Mean 37 medical mask Mean 33 N95 non fit tested Mean 35.3 N95 fit tested	Hospitals	4+1 weeks 4 weeks of intervention, 1 week non-wearing for development of respiratory symptoms.	- **Medical masks:** 3M™ medical mask, catalogue number 1820, St Paul, MN, USA - **N95 fit-tested mask:** 3M™ flat-fold N95 respirator, catalogue number 9132 - **N95 non-fit-tested mask:** 3M™ flat-fold N95 respirator, catalogue number 9132 - **Convenience control group** with no instructed mask use Selected from other hospitals where mask wearing was not routine during the study period. Absence of randomization in the no-mask group.	- **Clinical respiratory illness** Two or more respiratory or one respiratory symptom and a systemic symptom; - **ILI**, defined as fever ≥38°C plus one respiratory symptom (i.e. cough, runny nose, etc.) - **Laboratory-confirmed** viral respiratory infection	**Compliance** - 68% non-fit N95 - 74% fitN95 - 76 % in medical masks	- The rates of all outcomes were higher in The convenience no-mask group compared to The intervention arms. - There was no significant difference in outcomes between The N95 arms with and without fit testing.

MacIntyre et al., 2015	2011	Healthcare workers	Hanoi, Vietnam	**Total N = 1607 healthcare workers in 15 hospitals** - control group (N=458) - cloth mask (N=569) - medical mask (N=580) - Adults	Over 18 years of age Mean 36, CI(35.6 to 37.3)	Hospital wards	4+1 weeks 4 weeks of intervention, 1 week non-wearing for development of respiratory symptoms.	- **Cloth mask:** Cloth mask was provided - **Medical mask:** Medical mask was provided - **Control group:** No mask	**Clinical respiratory illness (CRI):** Defined as two or more respiratory symptoms or one respiratory symptom and a systemic symptom Influenza-like illness (ILI) defined as fever ≥38°C plus one respiratory symptom and Laboratory-confirmed viral respiratory infection. **Compliance with mask use:** Using the mask during the shift for 70% or more of work shift hours. - Participants were categorised as compliant if the average use was equal or more than 70% of the working time. Confounding factors were collected at baseline.	**Compliance over 70% of the time** - 56.6% of medical mask group - 56.8% of the clotch mask - 23.6% of the control group	- The rate of ILI was higher in the cloth masks arm compared with the control and medical mask group
MacIntyre et al., 2016	2013-2014	Household	Beijing, China	**Total N = 245 index cases and 597 contacts** - 122 index cases and 295 contacts in control arm - 123 index cases 302 contacts in mask arm - Adults	Individuals over 18 years of age Age mean 40.2 CI (37.6 to 42.8)	Household	7 days	- **Mask group:** surgical mask 3M 1817 - **Controls:** No mask	**Fever ≥38°C plus one respiratory symptom including cough, nasal congestion, runny nose, sore throat or sneezes.** Exclusion: no history of ILI among household members in the prior 14 days and who lived with at least two other people at home were reccruited for the study.	**Compliance** 35% controls used facemasks Post hoc analysis by mask use was conducted to account for compliance bias	- Intention to treat analysis did not show statistically significant protection with facemasks. - Facemask use vs. no mask use showed association with clinical infection but not with ILI.
Simmerman et al., 2011	2008-2009	Household	Bankok, Thailand	**Total N = 442 index children and 1147 household members.** - 155 index patient families 302 family members in control group (5 families dropped out). 155 index patient families, 292 family members in hand wash HW (8 families dropped out). 155 index patient families 291 family members in Hand wash plus mask (10 families dropped out). Index patients: children	Index patient age range from 0 to 15 years	Household	Up to 21 days	- **Control group:** nutritional, physical activity, and smoking cessation education - **Handwashing group:** Handwashing education and a handwashing kit, graduated dispenser with standard unscented liquid hand soap - **Handwashing with facemask:** households received handwashing education and the handwashing kit, and a box of 50 standard paper surgical face masks and 20 pediatric face masks: Med-con company, Thailand #14IN- 20AMB-30IN	**ILI was defined as fever >38°C and cough or sore throat in the absence of another explanation.** - illness <48 hours before respiratory specimens tested positive for influenza by a rapid influenza diagnostic test (RIDT) that was later confirmed by qualitative real-time RT-PCR (rRT-PCR). Exclusion: - Children at high risk for severe influenza complications and those treated with influenza antiviral medications - Households with any member reporting an ILI that preceded the index case by 7 days or less and households where any member had received influenza vaccination during the preceding 12 months	**Compliance** **- Index patients used masks 35minutes per day** - Parents wore their masks for a median of 153 minutes per day - siblings 17 minutes per day **Controls** - 17% control family members used facemasks	- No significant association with hand washing or with facemask with handwashing.
Suess et al., 2012	2009-2010, 2010-2011	Household	Berlin, Germany	**Total N = 218 individuals in 84 households** - 30 control house holds and 82 contacts - 26 Mask only house holds , 69 contacts - 28 Mask with hand hygiene house holds, 67 contacts - Children as index patients	Index cases were predominantly children under 14 years of age 92-100% per study arm. Contacts were predominantly adults median age control = 35, mask arm = 37, and mask and hand hygiene arm 34 years. During year 2009/2010 median age control = 38, mask arm = 35, and mask and hand hygiene arm 35 years.	Household	2 to 8 days	- **Mask and Hygiene arm:** alcohol based hand-rub: SteriliumTM, Bode Chemie, Germany Surgical facemasks in two different sizes, one for children aged younger than 14 years (Child's Face Mask, Kimberly-Clark, USA) and one for adults (Aérokyn Masques, LCH Medical Products, France) If masks intended for participants younger than 14 years did not fit properly request to wear adult masks instead Proper use of the interventions; - **Mask arm:** surgical facemasks and information on their correct use - **Control:** no masks or hand rub	**Included households with an influenza positive index case** in the absence of further respiratory illness within the preceding 14 days. Main outcome measure was laboratory confirmed influenza infection in a household contact. **A symptomatic secondary influenza virus infection definition:** - Laboratory confirmed influenza infection in a household member who developed fever (> 38.0°C), cough, or sore throat during the observation period. All other secondary cases were termed as subclinical. - A secondary outcome measure was the occurrence of ILI as defined as fever plus cough or sore throat.	**Adherence measured using surveys.** **Definition 1.** Daily adherence (always or mostly) **Definition 2.** Evaluation of the behavior during the first 5 days after implementation of the intervention (always or mostly). Daily compliance was over 50% in both M and MH. A gradual decline lower lower adherence began around the sixth day of the index patient's illness. Definition for the adherence for hand hygiene: if they had disinfected their hands at least five times per day on each of the 5 days **Mask and Hand hygiene group:** low hand hygiene adherence compared to the index patients. 5-day adherence in index patients of the MH group dropped from 41% to 18%. **Mask group:** Relatively stable adherence.	- Household transmission was reduced when implemented within 36 hours - Household transmission was not reduced when implemented after 36 hours - Per protocol analysis showed significant reduction of infections

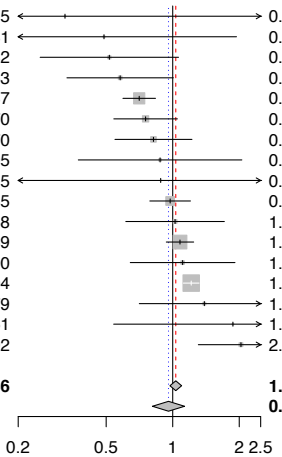
Table 1. Descriptive table outlining study characteristics in this systematic review and meta-analysis

Figure 3

A

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95%–CI	Weight (fixed)	Weight (random)
MacIntyre et al., 2016	1	302	3	295	0.326	[0.034; 3.112]	0.2%	0.5%
MacIntyre et al., 2011	3	492	6	481	0.489	[0.123; 1.943]	0.4%	1.3%
Suess et al., 2012	12	136	14	82	0.517	[0.251; 1.062]	1.2%	3.8%
Barasheed et al., 2014	11	36	28	53	0.578	[0.332; 1.007]	1.6%	5.3%
Aiello et al., 2010	191	745	177	487	0.705	[0.595; 0.836]	15.3%	11.3%
Aiello et al., 2012	77	741	51	370	0.754	[0.541; 1.050]	4.8%	8.5%
Bundgaard et al., 2020	42	2392	53	2470	0.818	[0.548; 1.222]	3.7%	7.4%
Cowling et al., 2008	6	61	23	205	0.877	[0.374; 2.055]	0.8%	3.0%
Jacobs et al., 2009	1	17	1	15	0.882	[0.060; 12.914]	0.1%	0.4%
Abdin et al., 2005	129	510	126	485	0.974	[0.788; 1.203]	9.2%	10.6%
Canini et al., 2010	24	148	25	158	1.025	[0.613; 1.712]	1.7%	5.8%
Alfelali et al., 2020	354	3199	322	3139	1.079	[0.935; 1.244]	23.2%	11.7%
MacIntyre et al., 2009	33	186	16	100	1.109	[0.643; 1.913]	1.5%	5.4%
Larson et al., 2010	575	938	457	904	1.213	[1.117; 1.316]	33.2%	12.4%
Cowling et al., 2009	18	258	14	279	1.390	[0.706; 2.738]	1.0%	4.1%
MacIntyre et al., 2015	14	1149	3	461	1.872	[0.541; 6.485]	0.3%	1.6%
Simmerman, et al., 2011	51	291	26	302	2.036	[1.306; 3.174]	1.8%	6.7%
Fixed effect model		**11601**		**10286**	**1.032**	**[0.972; 1.097]**	**100.0%**	**–**
Random effects model					**0.957**	**[0.810; 1.131]**	**–**	**100.0%**

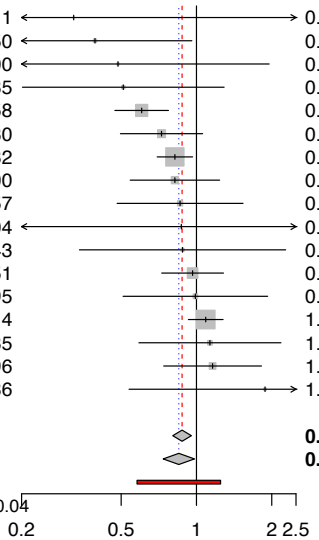
Heterogeneity: $I^2 = 73\%$, $\tau^2 = 0.0569$, $p < 0.01$



B

Study	TE	seTE	OR	95%–CI	Weight (fixed)	Weight (random)
MacIntyre et al., 2016	−1.130	1.1611	0.323	[0.033; 3.144]	0.1%	0.4%
Barasheed et al., 2014	−0.934	0.4550	0.393	[0.161; 0.959]	0.9%	2.3%
MacIntyre et al., 2011	−0.722	0.7100	0.486	[0.121; 1.953]	0.4%	1.0%
Suess et al., 2012	−0.673	0.4735	0.510	[0.202; 1.290]	0.9%	2.1%
Aiello et al., 2010	−0.504	0.1258	0.604	[0.472; 0.773]	12.1%	12.5%
Aiello et al., 2012	−0.322	0.1930	0.725	[0.497; 1.058]	5.1%	8.4%
Larson et al., 2010	−0.198	0.0832	0.820	[0.697; 0.965]	27.6%	15.8%
Bundgaard et al., 2020	−0.198	0.2100	0.820	[0.543; 1.238]	4.3%	7.6%
Cowling et al., 2009	−0.151	0.2957	0.860	[0.482; 1.535]	2.2%	4.7%
Jacobs et al., 2009	−0.139	1.4604	0.870	[0.050; 15.228]	0.1%	0.2%
Cowling et al., 2008	−0.128	0.4843	0.880	[0.341; 2.274]	0.8%	2.1%
Abdin et al., 2005	−0.036	0.1451	0.965	[0.726; 1.282]	9.1%	11.2%
Canini et al., 2010	−0.010	0.3395	0.990	[0.509; 1.926]	1.7%	3.8%
Alfelali et al., 2020	0.085	0.0814	1.089	[0.928; 1.277]	28.8%	16.0%
MacIntyre et al., 2009	0.124	0.3335	1.132	[0.589; 2.176]	1.7%	3.9%
Simmerman, et al., 2011	0.148	0.2296	1.160	[0.740; 1.819]	3.6%	6.8%
MacIntyre et al., 2015	0.631	0.6386	1.880	[0.538; 6.573]	0.5%	1.2%
Fixed effect model			**0.876**	**[0.804; 0.954]**	**100.0%**	**–**
Random effects model			**0.850**	**[0.736; 0.982]**	**–**	**100.0%**
Prediction interval				**[0.579; 1.249]**		

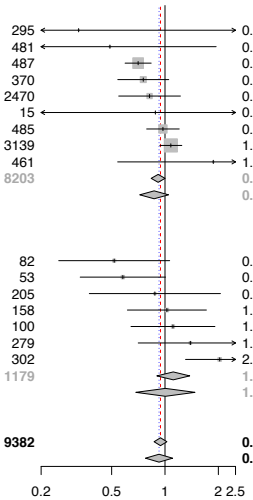
Heterogeneity: $I^2 = 42\%$, $\tau^2 = 0.0271$, $p = 0.04$



C

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95%–CI	Weight (fixed)	Weight (random)
Wearer protection								
Wearer protection								
MacIntyre et al., 2016	1	302	3	295	0.326	[0.034; 3.112]	0.3%	0.6%
MacIntyre et al., 2011	3	492	6	481	0.489	[0.123; 1.943]	0.6%	1.5%
Aiello et al., 2010	191	745	177	487	0.705	[0.595; 0.836]	22.8%	13.1%
Aiello et al., 2012	77	741	51	370	0.754	[0.541; 1.050]	7.3%	9.8%
Bundgaard et al., 2020	42	2392	53	2470	0.818	[0.548; 1.222]	5.6%	8.4%
Jacobs et al., 2009	1	17	1	15	0.882	[0.060; 12.914]	0.1%	0.4%
Abdin et al., 2005	129	510	126	485	0.974	[0.788; 1.203]	13.8%	12.3%
Alfelali et al., 2020	354	3199	322	3139	1.079	[0.935; 1.244]	34.7%	13.6%
MacIntyre et al., 2015	14	1149	3	461	1.872	[0.541; 6.485]	0.5%	1.8%
Fixed effect model		9547		8203	0.915	[0.835; 1.001]	85.6%	—
Source control					0.868	[0.720; 1.047]	—	61.3%
Heterogeneity: $I^2 = 58\%$, $\tau^2 = 0.0322$, $p = 0.02$								
Source control								
Suess et al., 2012	12	136	14	82	0.517	[0.251; 1.062]	1.9%	4.3%
Barasheed et al., 2014	11	36	28	53	0.578	[0.332; 1.007]	2.4%	6.0%
Cowling et al., 2008	6	61	23	205	0.877	[0.374; 2.055]	1.1%	3.3%
Canini et al., 2010	24	148	25	158	1.025	[0.613; 1.712]	2.6%	6.6%
MacIntyre et al., 2009	33	186	16	100	1.109	[0.643; 1.913]	2.2%	6.1%
Cowling et al., 2009	18	258	14	279	1.390	[0.706; 2.738]	1.4%	4.7%
Simmerman, et al., 2011	51	291	26	302	2.036	[1.306; 3.174]	2.7%	7.7%
Fixed effect model		1116		1179	1.113	[0.899; 1.379]	14.4%	—
Random effects model					1.005	[0.684; 1.477]	—	38.7%
Heterogeneity: $I^2 = 66\%$, $\tau^2 = 0.1729$, $p < 0.01$								
Fixed effect model		**10663**		**9382**	**0.943**	**[0.868; 1.025]**	**100.0%**	**–**
Random effects model					**0.926**	**[0.776; 1.104]**	**–**	**100.0%**

Heterogeneity: $I^2 = 61\%$, $\tau^2 = 0.0544$, $p < 0.01$
Residual heterogeneity: $I^2 = 62\%$, $p < 0.01$



D

Study	TE	seTE	OR	95%–CI	Weight (fixed)	Weight (random)
Wearer protection						
MacIntyre et al., 2016	−1.130	1.1611	0.323	[0.033; 3.144]	0.1%	0.4%
MacIntyre et al., 2011	−0.722	0.7100	0.486	[0.121; 1.953]	0.4%	1.0%
Aiello et al., 2010	−0.504	0.1258	0.604	[0.472; 0.773]	12.1%	12.5%
Aiello et al., 2012	−0.322	0.1930	0.725	[0.497; 1.058]	5.1%	8.4%
Bundgaard et al., 2020	−0.198	0.2100	0.820	[0.543; 1.238]	4.3%	7.6%
Jacobs et al., 2009	−0.139	1.4604	0.870	[0.050; 15.228]	0.1%	0.2%
Abdin et al., 2005	−0.036	0.1451	0.965	[0.726; 1.282]	9.1%	11.2%
Alfelali et al., 2020	0.085	0.0814	1.089	[0.928; 1.277]	28.8%	16.0%
MacIntyre et al., 2015	0.631	0.6386	1.880	[0.538; 6.573]	0.5%	1.2%
Fixed effect model			0.896	[0.803; 1.001]	60.6%	—
Random effects model			0.836	[0.661; 1.056]	—	58.5%
Heterogeneity: $I^2 = 60\%$, $\tau^2 = 0.0546$, $p = 0.01$						
Source control						
Barasheed et al., 2014	−0.934	0.4550	0.393	[0.161; 0.959]	0.9%	2.3%
Suess et al., 2012	−0.673	0.4735	0.510	[0.202; 1.290]	0.9%	2.1%
Larson et al., 2010	−0.198	0.0832	0.820	[0.697; 0.965]	27.6%	15.8%
Cowling et al., 2009	−0.151	0.2957	0.860	[0.482; 1.535]	2.2%	4.7%
Cowling et al., 2008	−0.128	0.4843	0.880	[0.341; 2.274]	0.8%	2.1%
Canini et al., 2010	−0.010	0.3395	0.990	[0.509; 1.926]	1.7%	3.8%
MacIntyre et al., 2009	0.124	0.3335	1.132	[0.589; 2.176]	1.7%	3.9%
Simmerman, et al., 2011	0.148	0.2296	1.160	[0.740; 1.819]	3.6%	6.8%
Fixed effect model			0.845	[0.738; 0.969]	39.4%	—
Random effects model			0.845	[0.738; 0.969]	—	41.5%
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.43$						
Fixed effect model			**0.876**	**[0.804; 0.954]**	**100.0%**	**–**
Random effects model			**0.850**	**[0.736; 0.982]**	**–**	**100.0%**
Prediction interval				**[0.579; 1.249]**		

Heterogeneity: $I^2 = 42\%$, $\tau^2 = 0.0271$, $p = 0.04$
Residual heterogeneity: $I^2 = 45\%$, $p = 0.03$

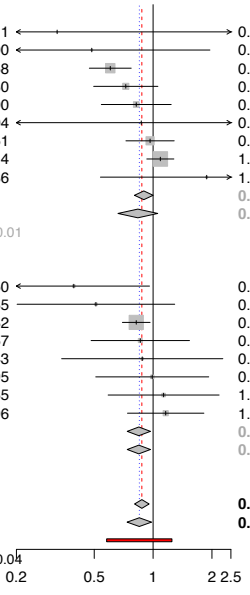


Figure 4

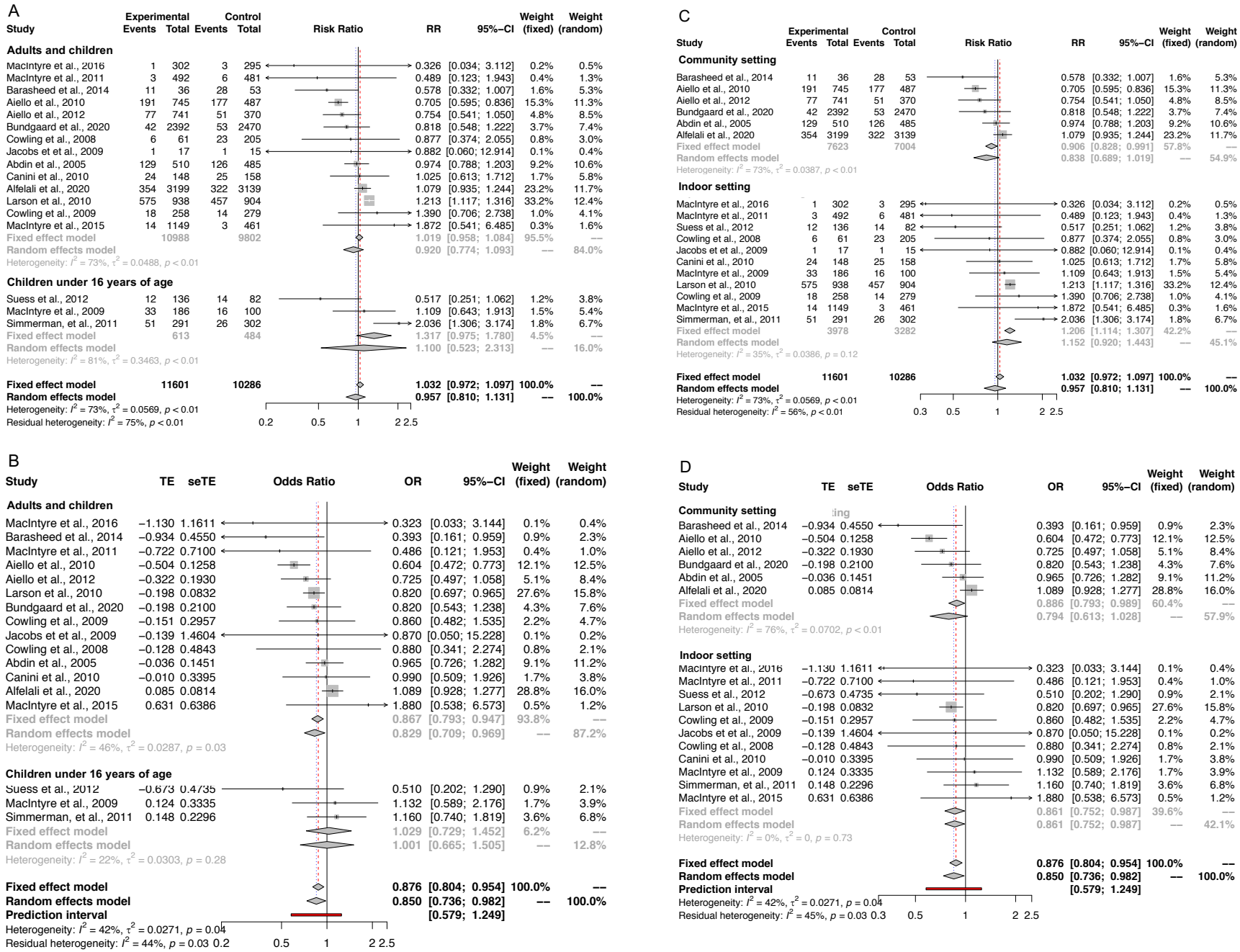
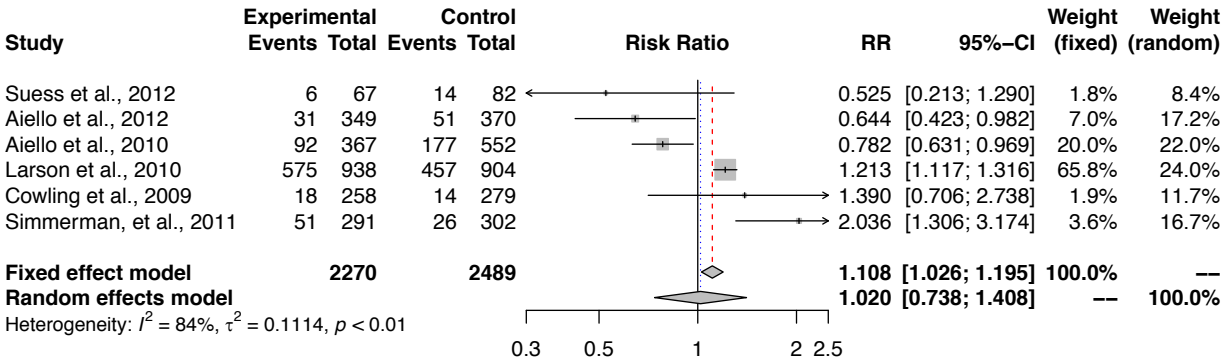
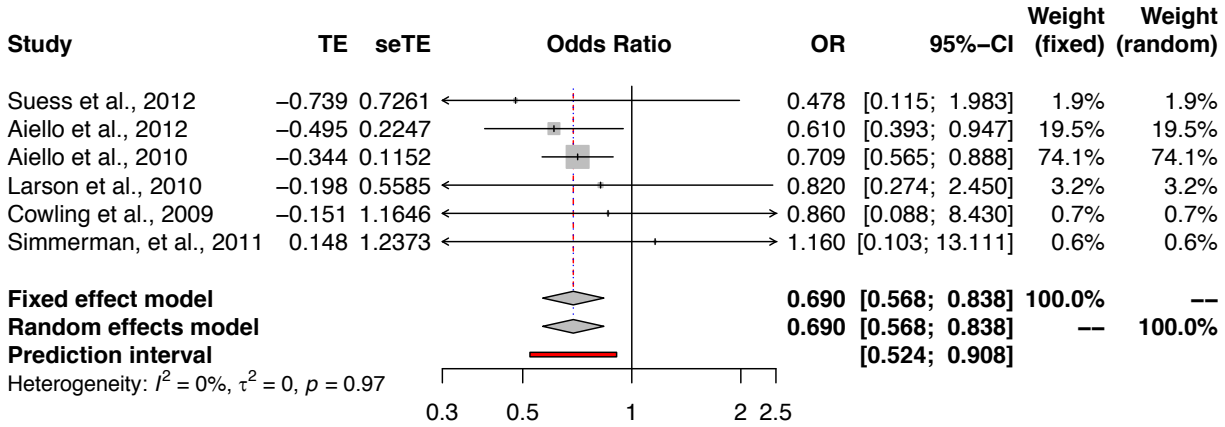

A
Study | Experimental Events Total | Control Events Total | Risk Ratio | RR 95%–CI | Weight (fixed) | Weight (random)
Adults and children
MacIntyre et al., 2016 1 302 3 295 0.326 [0.034; 3.112] 0.2% 0.5%
MacIntyre et al., 2011 3 492 6 481 0.489 [0.123; 1.943] 0.4% 1.3%
Barasheed et al., 2014 11 36 28 53 0.578 [0.332; 1.007] 1.6% 5.3%
Aiello et al., 2010 191 745 177 487 0.705 [0.595; 0.836] 15.3% 11.3%
Aiello et al., 2012 77 741 51 370 0.754 [0.541; 1.050] 4.8% 8.5%
Bundgaard et al., 2020 42 2392 53 2470 0.818 [0.548; 1.222] 3.7% 7.4%
Cowling et al., 2008 6 61 23 205 0.877 [0.374; 2.055] 0.8% 3.0%
Jacobs et al., 2009 1 17 1 15 0.882 [0.060; 12.914] 0.1% 0.4%
Abdin et al., 2005 129 510 126 485 0.974 [0.788; 1.203] 9.2% 10.6%
Canini et al., 2010 24 148 25 158 1.025 [0.613; 1.712] 1.7% 5.8%
Alfelali et al., 2020 354 3199 322 3139 1.079 [0.935; 1.244] 23.2% 11.7%
Larson et al., 2010 575 938 457 904 1.213 [1.117; 1.316] 33.2% 12.4%
Cowling et al., 2009 18 258 14 279 1.390 [0.706; 2.738] 1.0% 4.1%
MacIntyre et al., 2015 14 1149 3 461 1.872 [0.541; 6.485] 0.3% 1.6%
Fixed effect model 10988 9802 1.019 [0.958; 1.084] 95.5% —
Random effects model 0.920 [0.774; 1.093] — 84.0%
Heterogeneity: I^2 = 73%, τ^2 = 0.0488, p < 0.01
Children under 16 years of age
Suess et al., 2012 12 136 14 82 0.517 [0.251; 1.062] 1.2% 3.8%
MacIntyre et al., 2009 33 186 16 100 1.109 [0.643; 1.913] 1.5% 5.4%
Simmerman, et al., 2011 51 291 26 302 2.036 [1.306; 3.174] 1.8% 6.7%
Fixed effect model 613 484 1.317 [0.975; 1.780] 4.5% —
Random effects model 1.100 [0.523; 2.313] — 16.0%
Heterogeneity: I^2 = 81%, τ^2 = 0.3463, p < 0.01
Fixed effect model 11601 10286 1.032 [0.972; 1.097] 100.0% —
Random effects model 0.957 [0.810; 1.131] — 100.0%
Heterogeneity: I^2 = 73%, τ^2 = 0.0569, p < 0.01
Residual heterogeneity: I^2 = 75%, p < 0.01
0.2 0.5 1 2 2.5
B
Study | TE seTE | Odds Ratio | OR 95%–CI | Weight (fixed) | Weight (random)
Adults and children
MacIntyre et al., 2016 −1.130 1.1611 0.323 [0.033; 3.144] 0.1% 0.4%
Barasheed et al., 2014 −0.934 0.4550 0.393 [0.161; 0.959] 0.9% 2.3%
MacIntyre et al., 2011 −0.722 0.7100 0.486 [0.121; 1.953] 0.4% 1.0%
Aiello et al., 2010 −0.504 0.1258 0.604 [0.472; 0.773] 12.1% 12.5%
Aiello et al., 2012 −0.322 0.1930 0.725 [0.497; 1.058] 5.1% 8.4%
Larson et al., 2010 −0.198 0.0832 0.820 [0.697; 0.965] 27.6% 15.8%
Bundgaard et al., 2020 −0.198 0.2100 0.820 [0.543; 1.238] 4.3% 7.6%
Cowling et al., 2009 −0.151 0.2957 0.860 [0.482; 1.535] 2.2% 4.7%
Jacobs et al., 2009 −0.139 1.4604 0.870 [0.050; 15.228] 0.1% 0.2%
Cowling et al., 2008 −0.128 0.4843 0.880 [0.341; 2.274] 0.8% 2.1%
Abdin et al., 2005 −0.036 0.1451 0.965 [0.726; 1.282] 9.1% 11.2%
Canini et al., 2010 −0.010 0.3395 0.990 [0.509; 1.926] 1.7% 3.8%
Alfelali et al., 2020 0.085 0.0814 1.089 [0.928; 1.277] 28.8% 16.0%
MacIntyre et al., 2015 0.631 0.6386 1.880 [0.538; 6.573] 0.5% 1.2%
Fixed effect model 0.867 [0.793; 0.947] 93.8% —
Random effects model 0.829 [0.709; 0.969] — 87.2%
Heterogeneity: I^2 = 46%, τ^2 = 0.0287, p = 0.03
Children under 16 years of age
Suess et al., 2012 −0.673 0.4735 0.510 [0.202; 1.290] 0.9% 2.1%
MacIntyre et al., 2009 0.124 0.3335 1.132 [0.589; 2.176] 1.7% 3.9%
Simmerman, et al., 2011 0.148 0.2296 1.160 [0.740; 1.819] 3.6% 6.8%
Fixed effect model 1.029 [0.729; 1.452] 6.2% —
Random effects model 1.001 [0.665; 1.505] — 12.8%
Heterogeneity: I^2 = 22%, τ^2 = 0.0303, p = 0.28
Fixed effect model 0.876 [0.804; 0.954] 100.0% —
Random effects model 0.850 [0.736; 0.982] — 100.0%
Prediction interval [0.579; 1.249]
Heterogeneity: I^2 = 42%, τ^2 = 0.0271, p = 0.04
Residual heterogeneity: I^2 = 44%, p = 0.03 0.2 0.5 1 2 2.5
C
Study | Experimental Events Total | Control Events Total | Risk Ratio | RR 95%–CI | Weight (fixed) | Weight (random)
Community setting
Barasheed et al., 2014 11 36 28 53 0.578 [0.332; 1.007] 1.6% 5.3%
Aiello et al., 2010 191 745 177 487 0.705 [0.595; 0.836] 15.3% 11.3%
Aiello et al., 2012 77 741 51 370 0.754 [0.541; 1.050] 4.8% 8.5%
Bundgaard et al., 2020 42 2392 53 2470 0.818 [0.548; 1.222] 3.7% 7.4%
Abdin et al., 2005 129 510 126 485 0.974 [0.788; 1.203] 9.2% 10.6%
Alfelali et al., 2020 354 3199 322 3139 1.079 [0.935; 1.244] 23.2% 11.7%
Fixed effect model 7623 7004 0.906 [0.828; 0.991] 57.8% —
Random effects model 0.838 [0.689; 1.019] — 54.9%
Heterogeneity: I^2 = 73%, τ^2 = 0.0387, p < 0.01
Indoor setting
MacIntyre et al., 2016 1 302 3 295 0.326 [0.034; 3.112] 0.2% 0.5%
MacIntyre et al., 2011 3 492 6 481 0.489 [0.123; 1.943] 0.4% 1.3%
Suess et al., 2012 12 136 14 82 0.517 [0.251; 1.062] 1.2% 3.8%
Cowling et al., 2008 6 61 23 205 0.877 [0.374; 2.055] 0.8% 3.0%
Jacobs et al., 2009 1 17 1 15 0.882 [0.060; 12.914] 0.1% 0.4%
Canini et al., 2010 24 148 25 158 1.025 [0.613; 1.712] 1.7% 5.8%
MacIntyre et al., 2009 33 186 16 100 1.109 [0.643; 1.913] 1.5% 5.4%
Larson et al., 2010 575 938 457 904 1.213 [1.117; 1.316] 33.2% 12.4%
Cowling et al., 2009 18 258 14 279 1.390 [0.706; 2.738] 1.0% 4.1%
MacIntyre et al., 2015 14 1149 3 461 1.872 [0.541; 6.485] 0.3% 1.6%
Simmerman, et al., 2011 51 291 26 302 2.036 [1.306; 3.174] 1.8% 6.7%
Fixed effect model 3978 3282 1.206 [1.114; 1.307] 42.2% —
Random effects model 1.152 [0.920; 1.443] — 45.1%
Heterogeneity: I^2 = 35%, τ^2 = 0.0386, p = 0.12
Fixed effect model 11601 10286 1.032 [0.972; 1.097] 100.0% —
Random effects model 0.957 [0.810; 1.131] — 100.0%
Heterogeneity: I^2 = 73%, τ^2 = 0.0569, p < 0.01
Residual heterogeneity: I^2 = 56%, p < 0.01
0.3 0.5 1 2 2.5
D
Study | TE seTE | Odds Ratio | OR 95%–CI | Weight (fixed) | Weight (random)
Community setting
Barasheed et al., 2014 −0.934 0.4550 0.393 [0.161; 0.959] 0.9% 2.3%
Aiello et al., 2010 −0.504 0.1258 0.604 [0.472; 0.773] 12.1% 12.5%
Aiello et al., 2012 −0.322 0.1930 0.725 [0.497; 1.058] 5.1% 8.4%
Bundgaard et al., 2020 −0.198 0.2100 0.820 [0.543; 1.238] 4.3% 7.6%
Abdin et al., 2005 −0.036 0.1451 0.965 [0.726; 1.282] 9.1% 11.2%
Alfelali et al., 2020 0.085 0.0814 1.089 [0.928; 1.277] 28.8% 16.0%
Fixed effect model 0.886 [0.793; 0.989] 60.4% —
Random effects model 0.794 [0.613; 1.028] — 57.9%
Heterogeneity: I^2 = 76%, τ^2 = 0.0702, p < 0.01
Indoor setting
MacIntyre et al., 2016 −1.130 1.1611 0.323 [0.033; 3.144] 0.1% 0.4%
MacIntyre et al., 2011 −0.722 0.7100 0.486 [0.121; 1.953] 0.4% 1.0%
Suess et al., 2012 −0.673 0.4735 0.510 [0.202; 1.290] 0.9% 2.1%
Larson et al., 2010 −0.198 0.0832 0.820 [0.697; 0.965] 27.6% 15.8%
Cowling et al., 2009 −0.151 0.2957 0.860 [0.482; 1.535] 2.2% 4.7%
Jacobs et al., 2009 −0.139 1.4604 0.870 [0.050; 15.228] 0.1% 0.2%
Cowling et al., 2008 −0.128 0.4843 0.880 [0.341; 2.274] 0.8% 2.1%
Canini et al., 2010 −0.010 0.3395 0.990 [0.509; 1.926] 1.7% 3.8%
MacIntyre et al., 2009 0.124 0.3335 1.132 [0.589; 2.176] 1.7% 3.9%
Simmerman, et al., 2011 0.148 0.2296 1.160 [0.740; 1.819] 3.6% 6.8%
MacIntyre et al., 2015 0.631 0.6386 1.880 [0.538; 6.573] 0.5% 1.2%
Fixed effect model 0.861 [0.752; 0.987] 39.6% —
Random effects model 0.861 [0.752; 0.987] — 42.1%
Heterogeneity: I^2 = 0%, τ^2 = 0, p = 0.73
Fixed effect model 0.876 [0.804; 0.954] 100.0% —
Random effects model 0.850 [0.736; 0.982] — 100.0%
Prediction interval [0.579; 1.249]
Heterogeneity: I^2 = 42%, τ^2 = 0.0271, p = 0.04
Residual heterogeneity: I^2 = 45%, p = 0.03 0.3 0.5 1 2 2.5

Figure 5

A


Study
Experimental Events Total
Control Events Total
Risk Ratio
RR
95%–CI
Weight (fixed)
Weight (random)
Suess et al., 2012 6 67 14 82 0.525 [0.213; 1.290] 1.8% 8.4%
Aiello et al., 2012 31 349 51 370 0.644 [0.423; 0.982] 7.0% 17.2%
Aiello et al., 2010 92 367 177 552 0.782 [0.631; 0.969] 20.0% 22.0%
Larson et al., 2010 575 938 457 904 1.213 [1.117; 1.316] 65.8% 24.0%
Cowling et al., 2009 18 258 14 279 1.390 [0.706; 2.738] 1.9% 11.7%
Simmerman, et al., 2011 51 291 26 302 2.036 [1.306; 3.174] 3.6% 16.7%
Fixed effect model 2270 2489 1.108 [1.026; 1.195] 100.0% --
Random effects model 1.020 [0.738; 1.408] -- 100.0%
Heterogeneity: I^2 = 84%, τ^2 = 0.1114, p < 0.01
0.3 0.5 1 2 2.5

B


Study
TE seTE
Odds Ratio
OR
95%–CI
Weight (fixed)
Weight (random)
Suess et al., 2012 –0.739 0.7261 0.478 [0.115; 1.983] 1.9% 1.9%
Aiello et al., 2012 –0.495 0.2247 0.610 [0.393; 0.947] 19.5% 19.5%
Aiello et al., 2010 –0.344 0.1152 0.709 [0.565; 0.888] 74.1% 74.1%
Larson et al., 2010 –0.198 0.5585 0.820 [0.274; 2.450] 3.2% 3.2%
Cowling et al., 2009 –0.151 1.1646 0.860 [0.088; 8.430] 0.7% 0.7%
Simmerman, et al., 2011 0.148 1.2373 1.160 [0.103; 13.111] 0.6% 0.6%
Fixed effect model 0.690 [0.568; 0.838] 100.0% --
Random effects model 0.690 [0.568; 0.838] -- 100.0%
Prediction interval [0.524; 0.908]
Heterogeneity: I^2 = 0%, τ^2 = 0, p = 0.97
0.3 0.5 1 2 2.5

PRISMA 2009 Checklist

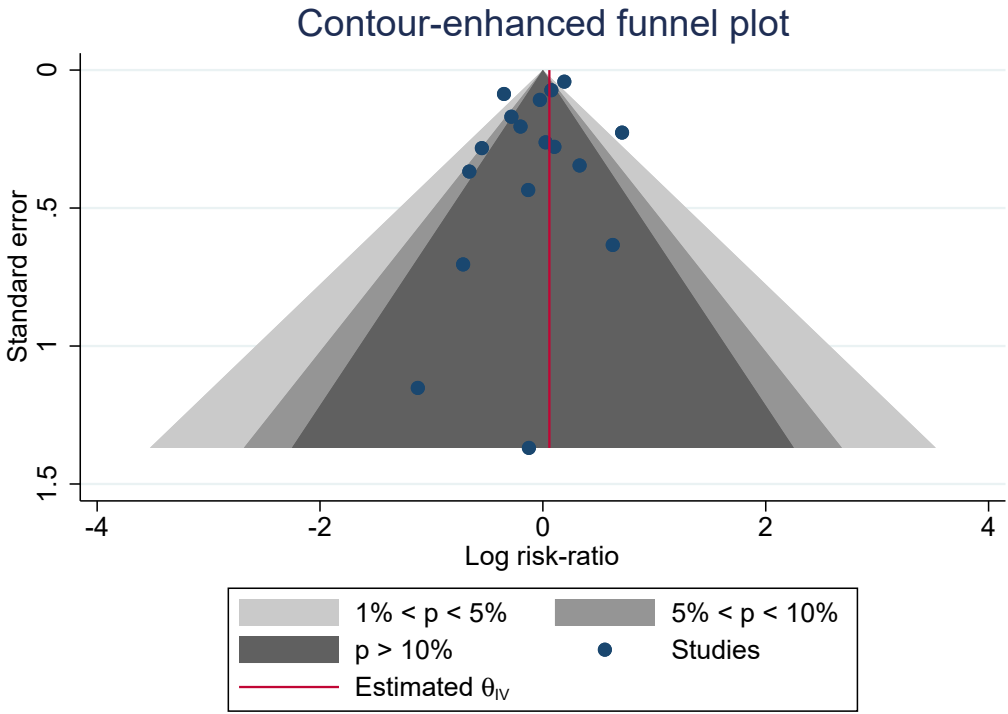
Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	cover
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	abstract
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	1
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	1
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	2
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	2
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	17-18
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	17-18
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	2-3
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	3
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	3
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	3
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	2
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	4

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	3-4
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	4
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	5
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Table1
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Fig 2 and 3
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Fig 2 and 3
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Fig 4
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Fig 2 and 3
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	8-10
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	10-11
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	9-10
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	11
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	cover and 11

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.


Contour-enhanced funnel plot
Standard error
0
.5
1
1.5
-4
-2
0
2
4
Log risk-ratio
1% < p < 5%
5% < p < 10%
p > 10%
Studies
Estimated θIV