

# A Mini-Review of the Low-Quality Evidence Underlying American Mask Policies



By now, we are all well familiar with the [fickle vacillation of public health guidance](#) around community masking. Initially, [masks were simply not really effective](#). Shortly thereafter, they were not only effective for protecting others, but [also for protecting oneself](#). Then they were [mandated](#).

Most recently, those cloth masks that have become commonplace, which were encouraged for nearly two years, which we were [taught to hand-make](#) by news outlets, were suddenly, as if overnight, relegated to '[facial decorations](#).'

How can it be that a tool which has been around and studied<sup>1</sup> for well over 100 years in the context of aerosolized respiratory viruses suddenly seems so poorly understood? This mini-review will advance the argument that low-quality evidence and impoverished bioethical frameworks have informed

a deeply fraught relationship to masking in America.

While we have surely all heard some variation of the argument that Americans are too myopic or selfish to do what people in [Asian countries](#) have been doing for decades, this is insufficient for making sense of the present moment. Ignoring the knowledge we have, avoiding cost-benefit analyses, and most importantly failing to clarify fundamental ethical principles risk irreversibly damaging the [credibility of medicine](#) and public health in the eyes of those we wish to serve.

## **Influenza-based studies of face mask efficacy**

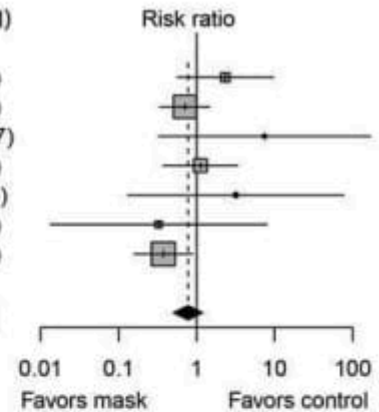
It is crucial to understand pre-COVID-19 research on mask efficacy in the context of influenza because, as was recognized early on, both respiratory pathogens are thought to be able to spread by breathing alone via exhaled aerosolized particles.<sup>2</sup> Prior to the COVID-19 pandemic, as late as 2019, the WHO Global Influenza Programme published an analysis of nonpharmaceutical interventions (NPIs) in the context of a potentially deadly viral respiratory pandemic,<sup>3</sup> at the time considered most likely to result from a novel influenza strain.

Culling systematic reviews of 18 NPIs including respiratory etiquette and face masks, the authors concluded that “[t]here is...a lack of evidence for the effectiveness of improved respiratory etiquette and the use of face masks in community settings during influenza epidemics and pandemics.” Nonetheless, the authors acknowledge that while “[t]here have been a number of high-quality randomized controlled trials (RCTs) demonstrating that personal protective measures such as hand hygiene and face masks have, at best, a small effect on influenza transmission, ... higher compliance in a severe pandemic might improve effectiveness.”

In early 2020, researchers in Hong Kong recognized the importance of diving into the pre-2020 literature on community masking more deeply. Acknowledging that “disposable medical masks...were designed to be worn by medical personnel to protect against accidental contamination of patient wounds, and to protect the wearer against splashes or sprays of bodily fluids,” University of Hong Kong researchers conducted a meta-analysis of the use of surgical masks to prevent influenza transmission in non-healthcare settings.<sup>4</sup> Their investigation concluded “[w]e did not find evidence that surgical-type face masks are effective in reducing laboratory-confirmed influenza transmission, either when worn by infected persons (source control) or by persons in the general community to reduce their susceptibility” (see Figure 1). These authors, much like the WHO authors, acknowledge in their discussion that masks might have value in reducing transmission of other infections when healthcare resources are stretched. However, that does not constitute positive evidence – it constitutes the absence of high-quality positive evidence.

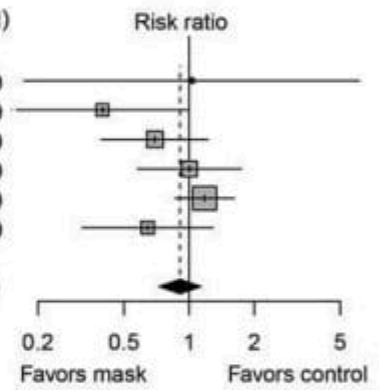
# A

Author (reference)	Mask		Control		Weight	Risk ratio (95% CI)
	Events	Total	Events	Total		
Aiello et al. 2010 (19)	5	347	3	487	5.7%	2.34 (0.56–9.72)
Aiello et al. 2012 (10)	12	392	16	370	37.3%	0.71 (0.34–1.48)
Barasheed et al. 2014 (33)	1	11	0	28	0.7%	7.43 (0.33–169.47)
Cowling et al. 2008 (12)	4	61	12	205	12.5%	1.12 (0.37–3.35)
MacIntyre et al. 2009 (34)	1	94	0	100	1.1%	3.19 (0.13–77.36)
MacIntyre et al. 2016 (35)	0	302	1	295	3.4%	0.33 (0.01–7.96)
Suess et al. 2012 (17)	6	69	19	82	39.4%	0.38 (0.16–0.89)
Fixed effect model	1,276		1,567		100.0%	0.78 (0.51–1.20)
Heterogeneity: $I^2 = 30\%$ , $\tau^2 = 0.1899$ , $p = 0.20$						
Test for overall effect: $z = -1.15$ ( $p = 0.25$ )						



# B

Author (reference)	Mask		Control		Weight	Risk ratio (95% CI)
	Events	Total	Events	Total		
Aiello et al. 2010 (9)	2	316	3	487	1.6%	1.03 (0.17–6.11)
Aiello et al. 2012 (10)	6	349	16	370	10.8%	0.40 (0.16–1.00)
Cowling et al. 2009 (11)	18	258	28	279	18.8%	0.70 (0.39–1.23)
Larson et al. 2010 (13)	25	938	24	904	17.1%	1.00 (0.58–1.74)
Simmerman et al. 2011 (15)	66	291	58	302	39.7%	1.18 (0.86–1.62)
Suess et al. 2012 (17)	10	67	19	82	11.9%	0.64 (0.32–1.29)
Fixed effect model	2,219		2,424		100.0%	0.91 (0.73–1.13)
Heterogeneity: $I^2 = 35\%$ , $\tau^2 = 0.0511$ , $p = 0.17$						
Test for overall effect: $z = -0.85$ ( $p = 0.39$ )						



# C

Author (reference)	Mask		Control		Weight	Risk ratio (95% CI)
	Events	Total	Events	Total		
Aiello et al. 2010 (9)	7	663	3	487	2.1%	1.71 (0.45–6.59)
Aiello et al. 2012 (10)	18	741	16	370	13.0%	0.56 (0.29–1.09)
Barasheed et al. 2014 (33)	1	11	0	28	0.2%	7.43 (0.33–169.47)
Cowling et al. 2009 (11)	18	258	28	279	16.3%	0.70 (0.39–1.23)
Cowling et al. 2008 (12)	4	61	12	205	3.3%	1.12 (0.37–3.35)
Larson et al. 2010 (13)	25	938	24	904	14.9%	1.00 (0.58–1.74)
MacIntyre et al. 2009 (34)	1	94	0	100	0.3%	3.19 (0.13–77.36)
MacIntyre et al. 2016 (35)	0	302	1	295	0.9%	0.33 (0.01–7.96)
Simmerman et al. 2011 (15)	66	291	58	302	34.6%	1.18 (0.86–1.62)
Suess et al. 2012 (17)	16	136	19	82	14.4%	0.51 (0.28–0.93)
Fixed effect model	3,495		3,052		100.0%	0.92 (0.75–1.12)
Heterogeneity: $I^2 = 30\%$ , $\tau^2 = 0.0593$ , $p = 0.17$						
Test for overall effect: $z = -0.84$ ( $p = 0.40$ )						

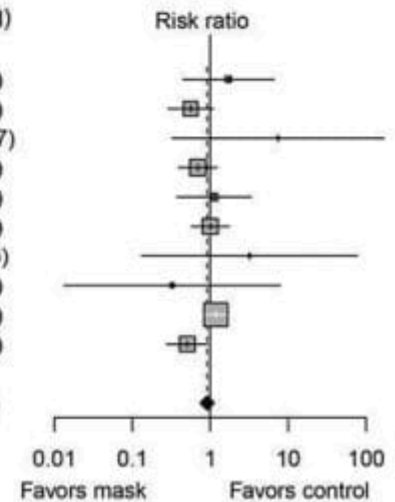


Figure 1: “Meta-analysis of risk ratios for the effect of face mask use with or without enhanced hand hygiene on laboratory-confirmed influenza from 10 randomized controlled trials with >6,500 participants. A) Face mask use alone; B) face mask and hand hygiene; C) face mask with or without hand hygiene. Pooled estimates were not made if there was high heterogeneity ( $I^2 > 75\%$ ). Squares indicate risk ratio for each of the included

studies, horizontal lines indicate 95% CIs, dashed vertical lines indicate pooled estimation of risk ratio, and diamonds indicate pooled estimation of risk ratio. Diamond width corresponds to the 95% CI.”<sup>4</sup>

In November of 2020, a Cochrane systematic review of 67 pre-pandemic RCTs and cluster-RCTs of physical interventions to reduce the spread of respiratory viruses was conducted.<sup>5</sup> Conclusions were striking:

*“The pooled results of randomized trials did not show a clear reduction in respiratory viral infection with the use of medical/surgical masks during seasonal influenza. There were no clear differences between the use of medical/surgical masks compared with N95/P2 respirators in healthcare workers when used in routine care to reduce respiratory viral infection. Hand hygiene is likely to modestly reduce the burden of respiratory illness. Harms associated with physical interventions were under-investigated.”*

Notably, this Cochrane review goes beyond community settings and raises questions about healthcare settings as well. When comparing surgical masks to no masks, the authors report moderate certainty evidence of little to no impact on the primary endpoint of laboratory-confirmed influenza based on a risk ratio of 0.91 in favor of masks, with a 95% confidence interval of 0.66 to 1.26.

Yet entering 2020, it was as if this body of literature never existed. So began the vigorous attempts to reinvent the wheel.

## **RCTs since the onset of the COVID-19 pandemic**

Throughout the pandemic, the Centers for Disease Control and Prevention has positioned itself as the authority on information about effective interventions to limit the spread

of the virus. Thus, their [web page](#) titled “Science Brief: Community Use of Masks to Control the Spread of SARS-CoV-2” is a natural resource from which to begin an investigation into the pandemic-era RCTs about masking.<sup>6</sup> Strikingly, there are only two RCTs discussed in any detail on this page. The first study cited on the page as supporting community masking is one of those RCTs – a “large, well-designed cluster-randomized trial in Bangladesh” conducted in late 2020. This is a widely circulated, well-regarded, properly controlled study and it makes sense why this would be listed first – it provides the strongest base of real-world, clinically relevant evidence for mask use in the setting of COVID-19 transmission.

What did the Bangladesh study show? After randomization of villages in rural Bangladesh to surgical mask, cloth mask, and no intervention arms, an intensive mask promotion strategy was carried out in intervention villages.<sup>7</sup> Researchers found the intervention led to a 29% absolute increase in proper mask-wearing in intervention villages. They also conclude that “[w]e find clear evidence that surgical masks lead to a relative reduction in symptomatic seroprevalence of 11.1% (adjusted prevalence ratio = 0.89 [0.78, 1.00]; control prevalence = 0.81%; treatment prevalence = 0.72%). Although the point estimates for cloth masks suggest that they reduce risk, the confidence limits include both an effect size similar to surgical masks and no effect at all.” In summary, the effects of cloth masks cannot be deemed statistically significant (no effect). Surgical masks, meanwhile, produced a 0.09% absolute risk reduction in symptomatic seropositivity relative to control. Converting this into a ‘number-needed-to-mask’ to prevent 1 instance of symptomatic seropositivity would come out to around 1,111 (1/0.0009). This number would be dramatically higher for endpoints of severe illness and death due to COVID-19.

What do these results mean? Of note, primary endpoints were not severe illness or death but rather having symptoms and

testing positive for COVID antibodies. Again, the authors reported a COVID seropositivity prevalence ratio (also known as risk ratio or relative risk) of 0.89 in the surgical mask vs no mask arms. In interpreting these results, we might compare them to the Cochrane review discussed above, finding a lab-confirmed influenza risk ratio of 0.91 in surgical mask vs no mask arms.

The Bangladesh results show minimally greater risk reduction in their mask arm when compared to this study. We might also compare the results to Figure 1 from the University of Hong Kong paper discussed above in which a lab-confirmed influenza risk ratio of 0.78 for surgical mask vs no mask was reported. The Bangladesh study shows a smaller effect in this comparison. Both of these influenza mask studies concluded surgical masks have basically no impact. All three of the studies discussed here had 95% confidence intervals including or crossing 1, the point at which surgical masks and no masks are associated with the same outcome. It would appear that prior to 2020, the effect size found by the Bangladesh study would be considered minimal at best and meaningless otherwise.

The second RCT on the CDC page is a study from Denmark.<sup>8</sup> These authors' priors (i.e. prior beliefs and expectations) revealed that they believed a 50% reduction in infection would be significant, and their study was conducted against this hypothesis. Priors are important because they shape what investigators are looking for. These authors did not find this reduction – instead they found an absolute risk reduction of 0.3% corresponding to a relative risk reduction of about 14% and a risk ratio of about 0.85 (95% confidence interval of about 0.72 to 0.99 per a letter to the editor).

Notably, the CDC concluded that the Bangladesh study showed that 'even modest increases in community use of masks can effectively reduce symptomatic SARS-CoV-2 infections.'<sup>6</sup> But

this raises many questions: What would it take to significantly increase effective use of masks in the community, above the 29% produced by the study? What would it do to the social fabric of a society to put that much effort into eliciting compliance to an intervention, all for a maximum absolute risk reduction for symptomatic seropositivity of less than 1% (again leaving aside endpoints of illness and death)? What does it mean that it took millions of dollars and a massive study in a foreign population with no baseline rate of vaccination to prove a small effect? And what does that suggest the impact of similar interventions might be in populations in this country?

## The state of the evidence

The questions raised above all point to another one – why were there not more RCTs to try to answer some of these questions? Many of the arguments for mask recommendations and mandates rest on [biological plausibility and filtration](#) studies, often relying on [mannequins](#). These simply cannot stand in for truly clinically-relevant data generated through large-scale randomized trials, especially when the force of public policy is being brought to bear through mask mandates. The real world is complicated. Factoring in real-world barriers to adherence is the only way to determine if an intervention is actually feasible and worth it. The evidence thus far consisting of very large systematic reviews, meta-analyses, and large RCTs does not appear to support such a policy.

As Dr. John P. Ioannidis has modeled, most published research findings in which investigators claim some relationship exists are likely to be false.<sup>9</sup> Many in the scientific community are also familiar with the reproducibility crisis within medical research. Thus, even if a new study were to come out claiming a much more significant effect size than those discussed above, it would need to be reproduced, and be subject to rigorous evaluation to assess for the latent biases which



Ioannidis identifies as undermining much of academic research.

In a January 2021 evidence review of masks in the setting of COVID-19, the authors propose some answers for why more RCTs have not been conducted.<sup>10</sup> “[E]thical issues,” they offer, “prevent the availability of an unmasked control arm.” They argue that “we should not generally expect to be able to find controlled trials, due to logistical and ethical reasons.” And yet it is precisely for ethical reasons that we must overcome the logistical hurdles to randomized controlled trials in order to prove efficacy.

Instead, we outsourced our ethical questions to rural villages in underdeveloped nations. If officials are going to expend political capital to bring the coercive power of the state to bear on enforcing behavior, at the bare minimum the evidence must be strong. But beyond this, a public debate about what the appropriate priors for further research and what the effect size should be to justify such action has not been held even two years into the pandemic. Both researchers and public health policymakers have failed to elucidate which bioethical principles they are operating from.

## **Ethical problems with justifying mandates**

Since mask mandates began getting implemented, mask-related policy has been driven by fallacious appeals to authority, reliance on low-quality evidence or minimal effect sizes, and violations of ethical principles like the precautionary principle and patient autonomy. The precautionary principle asserts that the burden is on those advocating intervention to prove the absence of harm and the definitive nature of the benefits. The principle of patient autonomy is central to medicine. Throughout the pandemic, the terrain on which masking rests has shifted. At times we have been told masking only protects oneself – at others we have been told that masking protects those nearby and thus it is imperative for

everyone to mask by a utilitarian ethic. In the 2020 Cochrane review, authors noted that harms were underinvestigated. This remains true.<sup>11</sup>

However, the problem with promoting low-quality evidence without grappling with fundamental ethical principles is that it leads to behavior and institutional decisions that may be completely out of touch with reality. For example, one's self-perception of risk may be inaccurate. An individual, overestimating the benefits of masking, might choose to visit a severely immunocompromised loved one believing they had eliminated most of the risk simply by masking. People might verbally or physically [attack](#) unmasked individuals with hostility out of a false belief that their risk of death is dramatically increased by the actions of others. A fear-ridden dermatologist wearing an N95 and face shield might ask an asymptomatic patient to hold his breath for the 5 seconds during which the mask was removed for a facial skin examination, believing it would meaningfully reduce her risk of COVID-19 infection. The [director of the CDC](#) might erroneously claim an absurdly high percentage, for example over 80%, by which 'masks' reduce one's chance of contracting COVID-19. And school districts in highly wealthy and educated enclaves [might transition children to wearing N-95s](#) despite the absence of validation studies in pediatric populations or community settings.

One might be compelled to ask: "What's the big deal? [#MaskLikeAKid!](#)" But these developments in our approach to infectious disease are not benign and are being implemented at scale. Human beings are being encouraged to view one another as perpetual vectors of disease and a relationship with the natural world based on resilience and harmony is being subordinated to a view of life as fundamentally dangerous, unsafe, and manageable with total control using methods for which we do not even have strong evidence.

While we can (and should) have vigorous debate about whether that view is appropriate in hospitals, it is surely inhumane to apply it to the rest of human life, especially in light of the fact that every respiratory pandemic has reached an inevitable state of endemicity.<sup>12</sup>

Medicine has a history of pathologizing the very things that connect us most to life on Earth from sunlight to our breath – this is not patient-centered, but anti-human. As [midterms approach](#), policy is beginning to change. But for two years, mask mandates were driven by the counterfactual question “What if many people die because we didn’t believe in masks enough?” This was no different from justifying enforced universal baptism by asking “What if many people go to Hell because we didn’t believe in God enough?” It is not science. It is Scientism.

## References

1. Kellogg WH, MacMillan G. An experimental study of the efficacy of gauze face masks. *American Journal of Public Health*. 1920;10(1):34-42.
2. Scheuch G. Breathing is enough: for the spread of influenza virus and SARS-CoV-2 by breathing only. *Journal of aerosol medicine and pulmonary drug delivery*. 2020;33(4):230-234.
3. Organization WH. *Non-pharmaceutical public health measures for mitigating the risk and impact of epidemic and pandemic influenza: annex: report of systematic literature reviews*. 2019.
4. Xiao J, Shiu EY, Gao H, et al. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—personal protective and environmental measures. *Emerging infectious diseases*. 2020;26(5):967.
5. Jefferson T, Del Mar CB, Dooley L, et al. Physical

interventions to interrupt or reduce the spread of respiratory viruses. *Cochrane database of systematic reviews*. 2020;(11)

6. Prevention CfDCa. Science Brief: Community Use of Masks to Control the Spread of SARS-CoV-2. Accessed February 4, 2022. <https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/masking-science-sars-cov2.html>

7. Abaluck J, Kwong LH, Styczynski A, et al. Impact of community masking on COVID-19: A cluster-randomized trial in Bangladesh. *Science*. 2021:eabi9069.

8. Bundgaard H, Ringgaard AK, Raaschou-Pedersen DET, Bundgaard JS, Iversen KK. Effectiveness of Adding a Mask Recommendation to Other Public Health Measures. *Annals of Internal Medicine*. 2021;174(8):1194-1195.

9. Ioannidis JP. Why most published research findings are false. *PLoS medicine*. 2005;2(8):e124.

10. Howard J, Huang A, Li Z, et al. An evidence review of face masks against COVID-19. *Proceedings of the National Academy of Sciences*. 2021;118(4)

11. Liu IT, Prasad V, Darrow JJ. How Effective Are Cloth Face Masks?: More than a Century after the 1918 Influenza Pandemic, Claims of the Masks' Effectiveness Continue to Lack a Firm Foundation. *Regulation*. 2021;44:32.

12. Heriot GS, Jamrozik E. Imagination and remembrance: what role should historical epidemiology play in a world bewitched by mathematical modelling of COVID-19 and other epidemics? *History and Philosophy of the Life Sciences*. 2021;43(2):1-5.