Unmasking the limitations of a recent COVID-19 masking study:

Did lifting school mask requirements really *cause* a higher COVID incident rate?

Ethan Moody, September 2023

Article Summary and Causal Claim

During the COVID-19 pandemic, public school masking policies emerged as a focal point of debate. Some argued that face-mask mandates enhanced student safety. Others viewed masking rules with skepticism. With public sentiment divided, the stage was set for scientific studies to offer much needed clarity on this subject.

One such study was conducted by Cowger et al. in 2022.[1] The aim of this study was to investigate the effect of universal masking policies on COVID incidence rates in public schools. The researchers examined weekly COVID case data from publicly available reports required by the Massachusetts Department of Elementary and Secondary Education (DESE) for students and staff across more than 70 public school districts in the greater Boston area. They used this data to statistically compare COVID case rates between two groups of districts: (1) those that lifted masking requirements either on or shortly after the recission of a statewide public school masking policy in February 2022 (70 districts), and (2) those that maintained these requirements for months after the policy change (only two districts, Boston and Chelsea). Citing the results of their analysis, the authors contend that "the lifting of masking requirements in school districts in the greater Boston area during March 2022 contributed an additional 45 Covid-19 cases per 1000 students and staff during the following 15-week period" (p. 1941) and notably impacted school attendance.

Cowger et al.'s primary claim is causal in nature: *the lifting of face-masking policies increased COVID incidence rates in certain school districts.* In both the abstract and discussion, the authors use causal language — e.g., "examine the *effect* of universal masking policies" (p. 1935, *emphasis added*) and "the lifting of masking requirements ... *contributed*" (p. 1941, *emphasis added*) — to describe the purpose of their study and its primary findings. They also include a short digression on the difference-in-differences technique and how it mitigates bias from confounders to strengthen causal inferences. This suggests that they really did intend to draw causal relationships from their work. Their conclusion that a universal masking policy affects COVID case rates and student/staff attendance is both compelling and highly practical in an educational setting, but does their study design actually allow them to make this claim?

Evaluation of Assumptions and Causal Effects

Although the authors attempt to eliminate confounding effects from sociodemographic characteristics and vaccination coverage across school districts, and though their data and timeframe support the parallel trend assumption for a difference-in-differences analysis, I'd argue that the study is not truly experimental. At best, it may be *quasi*-experimental. The study's design does not involve random assignment of subjects to treatment and control groups (a hallmark of experimental procedures) but instead leverages a group-split based on a "naturally-occurring event": the school districts' differing responses to the recission of a statewide COVID masking policy. As a result, the authors' claim about masking policy and its *causal effects* does not necessarily follow from the study; in fact, their study's conclusions are prone to bias. In order to believe their conclusion, we must assume that there are no other significant confounders influencing results, outside of the above factors. That is, we must believe that (1) there is no unobserved heterogeneity at play between the subjects in both school district groups (i.e., these subjects are essentially equivalent across all dimensions,

[1] Tori L. Cowger, Eleanor J. Murray, Jaylen Clarke, Marty T. Bassett, Bisola O. Ojikutu, Sarimer M. Sánchez, Natalia Linos, and Kathryn T. Hall. 2022. Lifting universal masking in schools — covid-19 incidence among students and staff. New England Journal of Medicine, 387(21):1935–1946. PMID: 36351262.

measured and unmeasured) — something which, in the absence of random assignment is not guaranteed — and (2) the baseline assumptions for difference-in-differences comparisons (like parallel trends in COVID case rates in the pre-intervention period) hold up.

These assumptions are not trivial, however. And given its design, the study is prone to bias. I'd argue that the purported causal effect is actually *overstated* as a result of several possible biasing factors. For example, if the students and staff in the 70 school districts that lifted masking requirements (the "lifted" group) reported COVID cases more frequently than those in the two school districts that maintained the masking requirements (the "maintained" group), this would have artificially inflated the effect from masking policy alone. Similarly, if those in the lifted group were biologically more susceptible to getting sick with COVID than those in the maintained group (e.g., because of immune system differences), the same would be true. Additionally, if those in the lifted group were already practicing risky behaviors that encouraged viral transmission or symptom emergence — like not washing hands, avoiding healthy food options, not exercising, and sleeping poorly — and these behaviors correlated with the mask-lifting decision, this would have made it seem as though the masking policy had a larger effect than it *actually* did. Furthermore, if the dominant COVID strain at the time of the comparison changed within some communities/school districts and not others, or if student family environments and/or staff support situations were different across the districts such that those in the maintained group had an easier time staying home from school when sick, it could be argued that these factors — not *just* the masking policy — impacted the observed differences in COVID rates across the district groups. Since the study does not control for these scenarios and relies on observational (non-experimental) data, there is a real possibility that the causal effects attributed to masking policy are both spurious and exaggerated.

Alternative Experiment

An experimental study could help to resolve these limitations. With public health at play, some experimental setups might encounter ethical challenges under frameworks like The Belmont Report, which advocate for beneficence and the principle of "do no harm" with respect to research subjects.[2] However, given the possibility of significant societal benefit resulting from learning more about the impact of COVID masking practices, the following experimental design could be feasible.

The experiment would focus on a single, large public school district — e.g., the Boston district — and randomly assign that district's students and staff members to one of two conditions for a period of 15 weeks: (1) maintain the statewide universal masking policy (control group), or (2) rescind the statewide universal masking policy (treatment group). Control group subjects would be required to wear face-masks at school in accordance with the policy, while treatment group subjects would not be required to do so. At the end of the 15-week period, COVID incidence rates (i.e., number of COVID-19 cases per 1000 students) would be compared between the two groups and evaluated for significant differences. If the treatment group showed a significantly higher incidence rate than the control group, the claim that lifting a universal masking policy (causally) increased school COVID rates would be supported. Otherwise, if the incidence rates were statistically equivalent across groups, this claim would be rejected. The focus on a single district would mitigate against geographic or other location-related confounders, and random assignment would prevent unobserved heterogeneity from influencing results. Admittedly, this design could suffer from fairness concerns (e.g., some students and staff would be placed in a potentially less favorable scenario than their peers/colleagues by not being required to wear masks) or political headwinds (e.g., the school district might not have sole power to enforce or administer disparate guidelines on individuals within its jurisdiction without the involvement of political authorities). Thoughtful considerations would also have to be made for families with multiple students enrolled, since those students could be randomly assigned to different conditions. However, clear participant consent procedures — such as the right to opt-out — and cooperation with political leaders and adherence to regulatory considerations could limit the severity of these issues as potential blockers for the study. The implementation of such an experiment would assist with the same benevolent goal as Cowger et al.'s study by shedding light on practices that support student and staff safety at school, while also providing the data and design components necessary to draw causal conclusions.

[2]Department of Health, Education, and Welfare; National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report. Ethical principles and guidelines for the protection of human subjects of research. J Am Coll Dent. 2014 Summer;81(3):4-13. PMID: 25951677.