UCSF
UC San Francisco Previously Published Works

Title

School Mask Mandates and COVID-19: The Challenge of Using Difference-in-Differences Analysis of Observational Data to Estimate the Effectiveness of a Public Health Intervention

Permalink

https://escholarship.org/uc/item/2ts2d4z4

Authors

Chandra, Ambarish
Hoeg, Tracy Beth
Ladhani, Shamez
et al.

Publication Date

2024-10-08

DOI

10.7326/M23-2907

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Annals of Internal Medicine MEDICINE AND PUBLIC ISSUES

School Mask Mandates and COVID-19: The Challenge of Using Difference-in-Differences Analysis of Observational Data to Estimate the Effectiveness of a Public Health Intervention

Ambarish Chandra, PhD*; Tracy Beth Høeg, MD, PhD*; Shamez Ladhani, MD, PhD; Vinay Prasad, MD, MPH; and Ram Duriseti, MD, PhD

Background: There are considerable challenges when using difference-in-differences (DiD) analysis of ecological data to estimate the effectiveness of public health interventions in rapidly changing situations.

Objective: To discuss the shortcomings of DiD methodology for the estimation of the effects of public health interventions using ecological data.

Design: As an example, the authors consider an analysis that used DiD methodology and reported a causal reduction in COVID-19 cases due to the maintenance of school mask mandates. They did alternate analyses using various control groups to assess the robustness of the prior analysis.

Setting: School districts in the greater Boston area and Massachusetts during the 2021-to-2022 academic year.

Participants: Students and school staff.

Measurements: Changes in COVID-19 case rates in districts that did and did not lift mask mandates.

Results: Important potential confounders rendered DiD methodology inappropriate for causal inference, including prior immunity, temporal variation in rates of infection, and changes in testing practices. The racial composition and income of intervention and control groups also differed substantially. Compared with maintaining the mask requirement, dropping the requirement was associated with anywhere from an increase of 5.64 cases (95% CI, 3.00 to 8.29 cases) per 1000 persons to a decrease of 2.74 cases (CI, 0.63 to 4.85 cases) per 1000 persons, depending on choice of control group and whether students or staff were examined.

Limitation: Ecological data were used; detailed data on all potential confounders were unavailable.

Conclusion: Alternate analyses yielded estimates consistent with a wide range of both negative and positive associations in COVID-19 case rates after removal of mask mandates. The findings highlight the challenges of using DiD analysis of ecological data to estimate the effectiveness of interventions in divergent intervention and control groups during rapidly changing circumstances.

Primary Funding Source: None.

Ann Intern Med. doi:10.7326/M23-2907 **Annals.org**
For author, article, and disclosure information, see end of text.
This article was published at Annals.org on 8 October 2024.
* Drs. Chandra and Høeg contributed equally to this work.

The use of masks for protection against COVID-19 in community settings has been highly controversial. Pooled analysis of randomized studies of medical masks to prevent the spread of respiratory viruses in 2020 and 2023 did not find evidence of benefit in the community (1, 2). One recent cluster randomized controlled trial reported marginal benefit of community surgical mask wearing, but follow-up analysis suggests this finding could be explained by sampling bias (3, 4). Observational studies of mask mandates in educational settings have had mixed findings (5–7).

Cowger and colleagues' ecological study of Boston-area school districts (8) used a staggered difference-in-differences (DiD) methodology to assess the effect of lifting mask mandates (February 2022) on COVID-19 incidence. In the greater Boston area, Chelsea and Boston were the only 2 districts that sustained masking requirements through June 2022 (Part C of the Supplement, available at Annals.org). The authors concluded that, among 72 school districts in the greater Boston area, lifting mask mandates resulted in an additional 44.9 cases (95% CI, 32.6 to 57.1 cases) per 1000 students and staff in the 15-week study period.

However, estimating causal effects from DiD methodology requires strong assumptions (9–12). Specifically, traditional DiD requires that, in the absence of an intervention, the unobserved difference in rates between the intervention and control groups would have remained constant ("parallel trends") and that no relevant factors change around the time of the intervention ("no confounding"). These conditions

See also:

Editorial comment
Related article

Web-Only
Supplement

Downloaded from https://annals.org by Univ of California San Francisco on 11/08/2024.

are unlikely to hold when evaluating the effect of COVID-19 interventions; indeed, methodologists experienced in these techniques have specifically cautioned against using DiD to evaluate the effect of COVID-19 policies (13).

The mask mandate intervention analysis by Cowger and colleagues contains multiple potential confounders that changed at the time of or after the change in mask mandates across school districts. These include differences in infection-based immunity, vaccination-based immunity, and SARS-CoV-2 testing practices. Districts that impose or drop mask mandates often simultaneously change contact tracing, case notifications, distancing, indoor lunch, music, physical education, and extracurricular practices, which could affect reported case rates (14–16). Some of these potential confounders were raised in letters to the editor about Cowger and colleagues' study (17–19).

Differences in race and socioeconomic status further increase the likelihood of confounding variables affecting postintervention trends, thus compromising the robustness of DiD methods for causal inference (11). Furthermore, limiting analyses to a single subset of districts and analyzing the results with a single method can produce spurious results (15, 20).

We reanalyzed the data from the original study by Cowger and colleagues (8) using alternative methods and control groups to explore the robustness of the conclusion about school mask mandates reducing COVID-19 case rates, and we used these findings to illustrate the assumptions and shortcomings of DiD methodology for the estimation of the effects of public health interventions from ecological data.

METHODS

We used the publicly available data of district case rates among students and staff in the 72 Boston-area school districts studied by Cowger and colleagues and added 3 alternative control groups in Massachusetts (Figure 1) (8, 21). We followed Cowger and colleagues by excluding districts that reported 0 cases for at least 10 weeks during the 2021-to-2022 school year (8). We analyzed differences in case rates for the academic year (1 September 2021 to 15 June 2022) among students and staff before and after mask mandate removal. Like Cowger and colleagues, we considered 3 March 2022, the first case-reporting date after the end of the statewide mandate, to be the change date of the mask policy.

Ethics

In accordance with 45 C.F.R. §46.102(f), this study was not submitted for institutional review board approval.

Data, Intervention, and Outcome

For each school district, data on weekly cases of COVID-19 and numbers of students and staff were publicly available from the Massachusetts Department of Elementary and Secondary Education (21). The primary exposure was the presence or absence of a masking requirement in each reporting week (8). The primary outcome was the weekly incidence of COVID-19 cases among students and staff after the change in school masking policies in March 2022. The change, if any, in masking mandates was unidirectional. Districts that dropped mandates did not subsequently reimpose them, with 2 exceptions noted in the Supplement.

Study Analysis and Extension

We recreated one of Cowger and colleagues' key figures, which plots district case rates for students from 1 January through 15 June 2022 by mask mandate policy and date of mask mandate removal, and then extended the figure to depict the entire academic year.

We defined 3 additional control groups: those within 50 km and 80 km of Boston, as well as the remainder of Massachusetts (Figure 1; Supplement Figures 1 and 2, available at Annals.org). Why Cowger and colleagues restricted their control group to school districts in the "New England city and town area of Boston–Cambridge–Newton" is unclear (8). First, the removal of mask mandates applied to all of Massachusetts (Supplement), and the Massachusetts Department of Public Health provided weekly data on testing and positive cases for the entire state. Second, multiple districts that lie outside Cowger and colleagues' control group are geographically closer to Boston and are socioeconomically more similar than other districts included in their study. Third, the borders of Cowger and colleagues' control group area do not correspond exactly with the "Boston–Cambridge–Newton" New England city and town area according to the U.S. Census Bureau (22).

Statistical Analysis

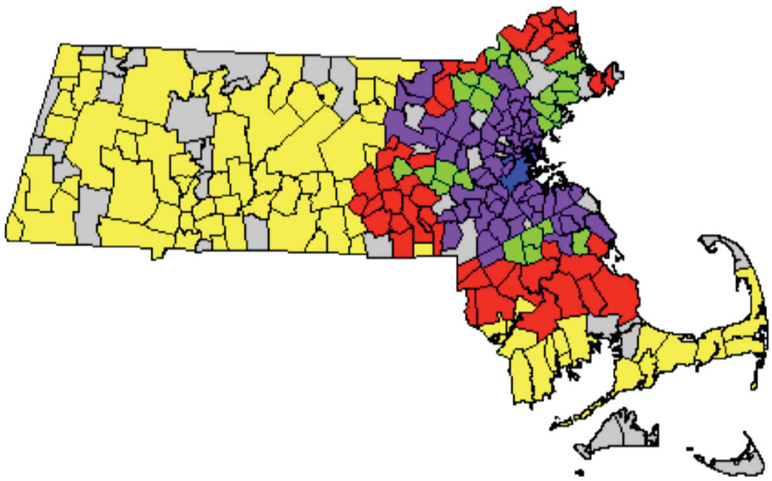
We calculated case rates and their absolute differences before and after 3 March 2022 for the intervention and control groups and stratified by students and staff using linear models. Linear model–based results correspond to reports in typical implementations of DiD analyses. We also calculated incidence rate ratios, and ratios of incidence rate ratios (RRRs), using a Poisson (log link) model, which easily allows for 0 case counts that arise in reported weekly case rates. The period-by-group interaction terms from these models estimate RRRs. A RRR greater than 1 indicates that districts that removed the masking requirement had a relative increase in the COVID-19 case rate compared with districts that maintained the requirement. Details about these calculations are provided in the Supplement. Raw data and code are posted publicly (23). In the Results section, we also point out where the findings are not consistent with the assumptions of a DiD analysis.

Role of the Funding Source

This study was not funded.

Downloaded from https://annals.org by Univ of California San Francisco on 11/08/2024.

Figure 1. School districts in Massachusetts, color-coded by study group, with corresponding socioeconomic information.



Color Key	Geographic Group	Districts, *n*	Median Income, *$*	Race: Percentage White
	Boston and Chelsea (maintained mask mandates)	2	51 600	10.7
	Cowger and colleagues' control districts	70	103 000	66.9
	<50 km* control districts	29	93 800	66.8
	<80 km* control districts	45	84 900	76.3
	Remainder of Massachusetts control	70	70 200	76.6

Insufficient data†

* Based on driving distance.
† Our analysis, like that of Cowger and colleagues (8), dropped districts that reported 0 cases over 10 weeks.

RESULTS

Figure 1 shows details about the intervention group (Boston and Chelsea) and various control groups. We identified a total of 215 school districts with available data across the state. The Boston and Chelsea districts are outliers according to race and income (**Supplement Figure 1**). Cowger and colleagues' control group had the highest median income. Comparing Boston and Chelsea with Cowger and colleagues' districts in a DiD analysis is problematic because, as described by Kahn-Lang and Lang (11), racial and socioeconomic differences between treatment and control groups tend to generate confounded postintervention results.

Analysis of Prior Study

Figure 2 shows weekly COVID-19 case rates per 1000 students from January to June 2022, with control groups stratified by mask mandate removal date (the corresponding figure for staff is **Supplement Figure 3**, available at Annals.org). Compared with Figure 1B in Cowger and colleagues' article, the full **Figure 2** extends the time period to the entire 2021-to-2022 academic year to show that the magnitude of differences in rates across districts was inconsistent over time (or did not visibly show parallel trends in those rates).

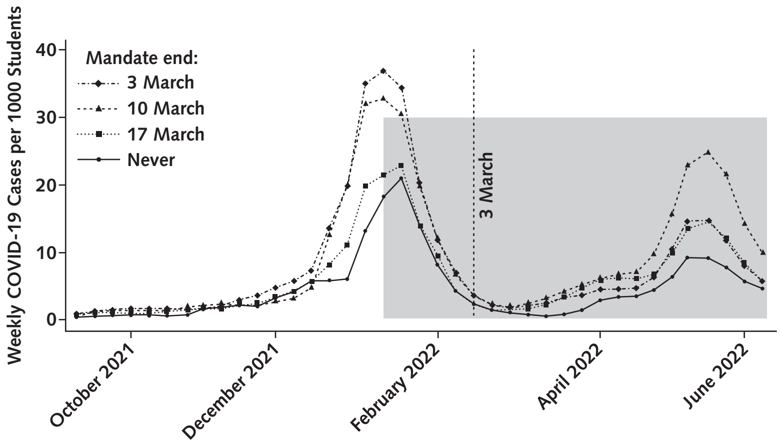
In addition, the largest decrease in cases from before to after the mandate (22%) occurred in the districts that dropped their mask mandates first (dotted-and-dashed line in **Figure 2**) and exceeded the corresponding 12% decline in Boston–Chelsea districts (solid line in **Figure 2**). This finding is inconsistent with a causal effect of changes in mask policy and does not support the use of a staggered treatment DiD analysis (8, 24).

Prior Community Cases

In contrast to case rates within schools, prior case rates in the community were highest in the towns of Boston and Chelsea (**Figure 3**). This discrepancy may be explained by different testing patterns in schools from those in communities (see the following section, Testing Rates). The community containing the Boston–Chelsea districts, which never ended school mask mandates, consistently reported *higher* cumulative case rates and exhibited the most rapid rise in case rates immediately preceding the end of the statewide mask mandate (**Figure 3**). This premandate surge in infections was likely a confounder, because these earlier infections decreased susceptibility to later infection in Boston and Chelsea after 3 March (25). In fact, we found that prior community case rates exhibited a

Downloaded from https://annals.org by Univ of California San Francisco on 11/08/2024.

Figure 2. COVID-19 case rates among students, September 2021–June 2022.



The shaded gray region denotes the period January–June 2022 shown in Figure 1B of Cowger and colleagues' article (8). The full figure corresponds to the entire 2021–2022 academic year. District groups are patterned by mask mandate drop date (or none).

substantial negative correlation (−0.60 [95% CI, −0.74 to −0.43]) with community case rates after the mask mandate change.

Testing Rates

According to a publication using internal data from the Massachusetts Department of Elementary and Secondary Education, consent to school SARS-CoV-2 testing in the 2021-to-2022 academic year was lower among Black students (<30%) than among White students (>60%) (26). This variation could confound reported case rates across school districts because Boston and Chelsea had a much higher fraction of Black students than almost all other districts (Figure 1). Transition to at-home testing may have had similar, larger, or smaller discrepancies in terms of participation by race or district, creating another potential source of confounding (27, 28).

Infectious Waves Robustness Check

An analysis of the 2 major Omicron waves during the winter before the mask policy change (15 December 2021 to 20 January 2022) and the spring after the mask policy change (28 April to 1 June 2022) comparing Boston and Chelsea with Cowger and colleagues' control districts showed that the ratio of cases in Boston and Chelsea to the unmasked districts was *higher* in spring (54.3% lower) than in winter (45.9% lower) after the policy change (Supplement). This is not compatible with attribution of causality to mask mandates for observed lower case rates using DiD.

Alternate Control Groups and Analyses

The Table presents changes in case rates by intervention and control groups (defined in Figure 1). Depending on the choice of control school districts, the dropping of a mask requirement was associated with anywhere from an increase of 5.64 cases (CI, 3.00 to 8.29 cases) per 1000 persons to a decrease of 2.74 cases (CI, 0.63 to 4.85 cases) per 1000 persons, a range that is consistent with either positive or negative associations between mask mandates and case rates. Similarly, RRRs ranged from 2.09 (CI, 1.76 to 2.48) among staff using Cowger and colleagues' control group to 0.61 (CI, 0.48 to 0.76) among students when the control group was districts less than 80 km away.

Geographic Robustness Check

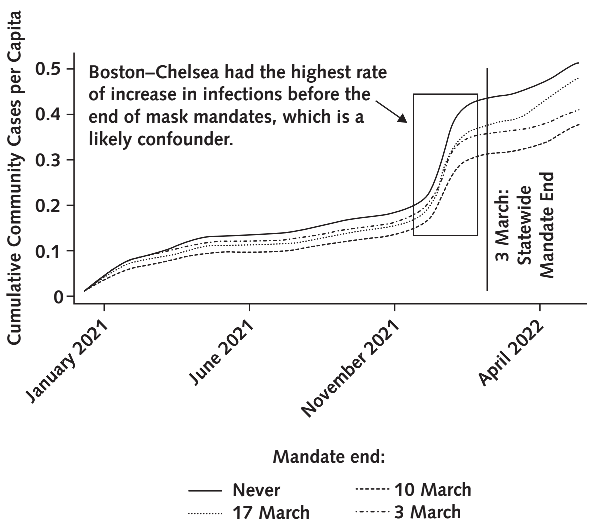
Examining all districts in Massachusetts, we found that student case rates were lower in Suffolk County (the Boston and Chelsea districts) than in all other counties statewide before 3 March 2022. This difference narrowed by 37% in the spring (Supplement Figure 2). Confounding from differential preintervention trends in prior infection–related immunity violates the assumptions required to justify causal inference from DiD.

Discussion

Our findings highlight the challenges of using DiD methods for the analysis of ecological data to estimate the effectiveness of COVID-19 policy interventions in intervention and control groups with widely different population characteristics during periods of rapidly changing infection rates. We show that the estimated effects of the intervention were not robust to changes in study design, such as the use of different control groups in statewide analysis, data restriction to only major infectious waves, or stratification by students and staff. In the present example, we found that maintaining school mask mandates was compatible with both significant decreases and increases in district case rates depending on the choice of control group and whether students or staff were studied. Our findings reinforce the

Downloaded from https://annals.org by Univ of California San Francisco on 11/08/2024.

Figure 3. Cumulative community cases, by district mask policy.



The figure plots cumulative community cases of COVID-19 per capita for the 4 groups of school districts studied by Cowger and colleagues (8), based on town-level data from the Massachusetts Department of Public Health.

pitfalls of using a limited set of observational data and DiD methodology to draw conclusions about intervention efficacy, particularly when the risk for confounding is high. In this circumstance, the DiD methodology for causal inference did not meet the required assumptions (9). Of note, our study identified multiple potential confounders, which may have differentially affected postintervention trends, including recent changes in prior community immunity and testing practices. Investigators have cautioned against the use of DiD for establishing the efficacy of pandemic mitigation strategies due to rapidly changing conditions and the accompanying challenges of meeting the assumptions required for causal inference (13).

The 2 districts that maintained mask mandates were clear demographic outliers. They were composed of only 11% White students, compared with 65% to 80% White across the remainder of the state (29). The control group in the initial study had a median income almost twice that of the Boston and Chelsea districts and substantially higher than that of the rest of the state. Large racial and socioeconomic differences between groups can be problematic for causal inference from DiD analyses (11).

Our study identified marked differences in recent community prior immunity, which could explain some portion of the case rate variation in spring 2022. We found that the higher the community case rates were before the week ending on 3 March, the lower the case rates were in the spring. Unlike district case rates, community case rates were highest in the Boston–Chelsea districts before 3 March and rose most dramatically going into the mask policy change period. This variation in preintervention case rates is likely an important confounder, violating a basic assumption of using the DiD methodology. We are not the first investigators to raise this issue with regard to the current example (17–19, 30).

Changes in testing practices across school districts and leading up to 3 March may act as another source of confounding. Socioeconomic characteristics and Massachusetts data suggest that at-home testing rates and consent to school testing (and therefore case reporting), respectively, may have been lower in the Boston–Chelsea districts (26–28).

Finally, the particular staggered DiD methodology used by Cowger and colleagues for causal inference is normally applied to policy implementations staggered over years and not weeks (24). We show that findings must meet the assumptions implicit in DiD methodology for causal inference.

Other statistical methods can be applied to evaluate natural experiments studies in school settings with similar intervention and control groups. Regression discontinuity or crossover designs can decrease the chances of confounding. Such designs have failed to identify a substantial benefit of mask mandates (6, 7, 31). One regression discontinuity study exploited a policy cutoff to compare unmasked 5-year-old students with masked 6-year-old students over the same time frame in the same setting, reducing the chance of confounding by differences in location, socioeconomic factors, or other policies changing with the mask policies. The crossover period in the second study allowed comparison of the 2 adjacent districts both while they had different mask policies and while they had the same policy, which could be used to help identify residual confounding variables between the districts. Thus, these particular designs were less likely to be confounded by disparities in prior infection, testing, or socioeconomic differences.

Some authors have argued that the lack of randomized controlled trials to provide high-quality evidence is itself a major pandemic failure (32–34). For example, cluster randomized controlled trials might be preferable for evaluating pandemic interventions. Randomized studies in kindergarten through 12th grade could be designed with classrooms as clusters. Target trial simulation matching classrooms with similar characteristics that did and did not have mask mandates might be considered; however, residual confounding may be difficult to avoid or rule out (35). In addition, because more than 95% of identified SARS-CoV-2 infections in schools have been traced to the surrounding community, school case rates will tend to reflect community rates, thereby both confounding and diluting any effect of school masking (36–39). In an ideal study, randomization would take place at the classroom or school level so as not to be biased by differences in surrounding communities with dissimilar policies, case rates, or underlying immunity rates. Even relatively small, well-

Downloaded from https://annals.org by Univ of California San Francisco on 11/08/2024.

Table. Case Rates and Relative Differences for Treatment and Various Control Groups

Variable	Weekly Cases per 1000 Persons in Group			
	Before Mask Policy Change	**After Mask Policy Change**	**Pre-Post Difference***	**RRR†**
Students				
Boston–Chelsea (mask mandates)	4.53	2.87	−1.66	–
Cowger and colleagues' control districts	8.98	7.78	−1.2	–
Difference (95% CI)	4.45	4.91	**0.46** (−1.64 to 2.57)	**1.05** (0.83 to 1.33)
Alternate control 2 (<50 km)	9.89	6.5	−3.39	–
Difference from Boston–Chelsea (95% CI)	5.36	3.63	**−1.73** (−4.06 to 0.59)	**0.61** (0.45 to 0.83)
Alternate control 3 (<80 km)	10.06	5.98	−4.08	–
Difference from Boston–Chelsea (95% CI)	5.53	3.11	**−2.42** (−4.42 to -0.42)	**0.61** (0.48 to 0.76)
Alternate control 4 (>80 km)	10.74	6.34	−4.4	–
Difference from Boston–Chelsea (95% CI)	6.21	3.47	**−2.74** (−4.85 to −0.63)	**0.61** (0.48 to 0.77)
Staff				
Boston–Chelsea	12.43	9.64	−2.79	–
Cowger and colleagues' control districts	9.84	12.69	2.85	–
Difference	−2.59	3.05	**5.64** (3.00 to 8.29)	**2.09** (1.76 to 2.48)
Alternate control 2 (<50 km)	10.92	11.19	0.27	–
Difference from Boston–Chelsea	−1.51	1.55	**3.06** (0.09 to 6.03)	**1.50** (1.21 to 1.85)
Alternate control 3 (<80 km)	11.55	11.22	−0.33	–
Difference from Boston–Chelsea	−0.88	1.58	**2.46** (−0.22 to 5.12)	**1.47** (1.24 to 1.74)
Alternate control 4 (>80 km)	11.35	12	0.65	–
Difference from Boston–Chelsea	−1.08	2.36	**3.44** (0.59 to 6.29)	**1.53** (1.25 to 1.87)

RRR = ratio of incidence rate ratios.
* Linear difference pre–post estimates are the bolded values in this column. Positive values indicate that removing mask requirements was associated with a relative increase in COVID-19 cases.
† RRRs are calculated using Poisson regressions and represent the ratio between the treatment and control groups of the ratio of case rates from before to after the intervention. A RRR >1 indicates that removing mask requirements was associated with a relative increase in COVID-19 cases. CIs for both calculations are derived from SEs that are clustered by district.

designed randomized trials may be able to rule out large effects, as was achieved with the DANMASK (Danish Study to Assess Face Masks for the Protection Against COVID-19 Infection) trial (40). Although a beneficial effect of any intervention may seem plausible, most do not end up being shown to be effective in randomized trials (33, 41).

In conclusion, our analyses show how observational studies of pandemic mitigation strategies using common DiD analysis of ecological data can produce noninformative or spurious results. Observational studies of public health interventions require extraordinary care in design and interpretation.

From Department of Management, University of Toronto Scarborough, and Rotman School of Management, University of Toronto, Toronto, Ontario, Canada (A.C.); Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts; Department of Emergency Medicine, University of California, San Francisco, San Francisco, California; and Department of Clinical Research, University of Southern Denmark, Odense, Denmark (T.B.H.); Paediatric Infectious Diseases Research Group, St. George's University of London, and Immunisation and Countermeasures Division, UK Health Security Agency, London, United Kingdom (S.L.); Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California (V.P.); and Department of Emergency Medicine, Stanford School of Medicine, Palo Alto, California (R.D.).

Disclosures: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M23-2907.

Corresponding Author: Tracy Beth Høeg, MD, PhD, Sloan School of Management, Massachusetts Institute of Technology, Office E62-562, 100 Main Street, Cambridge, MA 02142; e-mail, tbhoeg@mit.edu.

Author contributions are available at Annals.org.

References

1. **Jefferson T, Del Mar CB, Dooley L, et al.** Physical interventions to interrupt or reduce the spread of respiratory viruses. Cochrane Database Syst Rev. 2020;11:CD006207. [PMID: 33215698] doi:10.1002/14651858.CD006207.pub5
2. **Jefferson T, Dooley L, Ferroni E, et al.** Physical interventions to interrupt or reduce the spread of respiratory viruses. Cochrane Database Syst Rev. 2023;1:CD006207. [PMID: 36715243] doi:10.1002/14651858.CD006207.pub6
3. **Abaluck J, Kwong LH, Styczynski A, et al.** Impact of community masking on COVID-19: a cluster-randomized trial in Bangladesh. Science. 2022;375:eabi9069. [PMID: 34855513] doi:10.1126/science.abi9069
4. **Chikina M, Pegden W, Recht B.** Re-analysis on the statistical sampling biases of a mask promotion trial in Bangladesh: a statistical replication. Trials. 2022;23:786. [PMID: 36109816] doi:10.1186/s13063-022-06704-z
5. **Budzyn SE, Panaggio MJ, Parks SE, et al.** Pediatric COVID-19 cases in counties with and without school mask requirements – United States, July 1–September 4, 2021. MMWR Morb Mortal Wkly Rep. 2021;70:1377-1378. [PMID: 34591829] doi:10.15585/mmwr.mm7039e3

Downloaded from https://annals.org by Univ of California San Francisco on 11/08/2024.

6. **Sood N, Heick S, Stevenson J, et al.** Association between school mask mandates and SARS-CoV-2 student infections: evidence from a natural experiment of neighboring K-12 districts in North Dakota. Research Square. Preprint posted online 1 July 2022. doi:10.21203/rs.3.rs-1773983/v1
7. **Coma E, Català M, Méndez-Boo L, et al.** Unravelling the role of the mandatory use of face covering masks for the control of SARS-CoV-2 in schools: a quasi-experimental study nested in a population-based cohort in Catalonia (Spain). Arch Dis Child. 2023;108:131-136. [PMID: 35999036] doi:10.1136/archdischild-2022-324172
8. **Cowger TL, Murray EJ, Clarke J, et al.** Lifting universal masking in schools — Covid-19 incidence among students and staff. N Engl J Med. 2022;387:1935-1946. [PMID: 36351262] doi:10.1056/NEJMoa2211029
9. **Bertrand M, Duflo E, Mullainathan S.** How much should we trust differences-in-differences estimates? Q J Econ. 2004;119:249-275. [10.1162/003355304772839588]
10. **Wing C, Simon K, Bello-Gomez RA.** Designing difference in difference studies: best practices for public health policy research. Annu Rev Public Health. 2018;39:453-469. [PMID: 29328877] doi:10.1146/annurev-publhealth-040617-013507
11. **Kahn-Lang A, Lang K.** The promise and pitfalls of differences-in-differences: reflections on 16 and Pregnant and other applications. J Bus Econ Stat. 2020;38:613-620. [10.1080/07350015.2018.1546591]
12. **Zeldow B, Hatfield LA.** Confounding and regression adjustment in difference-in-differences studies. Health Serv Res. 2021;56:932-941. [PMID: 33978956] doi:10.1111/1475-6773.13666
13. **Callaway B, Li T.** Policy evaluation during a pandemic. J Econom. 2023;236:105454. [PMID: 37359750] doi:10.1016/j.jeconom.2023.03.009
14. **Høeg TB, Prasad V, Porter T.** Contact tracing policy for masked students may be an important confounding variable. Pediatrics. 2022;150:e2022057636A. doi:10.1542/peds.2022-057636A
15. **Chandra A, Høeg TB.** Lack of correlation between school mask mandates and paediatric COVID-19 cases in a large cohort. J Infect. 2022;85:671-675. [PMID: 36183909] doi:10.1016/j.jinf.2022.09.019
16. **Ladhani SN.** Face masking for children - time to reconsider. J Infect. 2022;85:623-624. [PMID: 36170896] doi:10.1016/j.jinf.2022.09.020
17. **Ladhani SN, Andrews N, Ramsay ME.** Lifting universal school masking — Covid-19 incidence among students and staff. N Engl J Med. 2023;389:579. [PMID: 37590461] doi:10.1056/NEJMc2215560
18. **Marine JE.** Lifting universal school masking — Covid-19 incidence among students and staff. N Engl J Med. 2023;389:579. [PMID: 37590462] doi:10.1056/NEJMc2215560
19. **Russell FM, Curtis N, Carlin J.** Lifting universal school masking — Covid-19 incidence among students and staff. N Engl J Med. 2023;389:578-579. [PMID: 37590460] doi:10.1056/NEJMc2215560
20. **Silberzahn R, Uhlmann EL, Martin DP, et al.** Many analysts, one data set: making transparent how variations in analytic choices affect results. Adv Methods Pract Psychol Sci. 2018;1:337-356. [10.1177/2515245917747646]
21. **Massachusetts Department of Elementary and Secondary Education.** Positive COVID-19 cases in schools. Accessed at https://archives.lib.state.ma.us/handle/2452/833423 on 3 January 2023.
22. **U.S. Census Bureau.** New England City and Town Areas: February, 2013. Accessed at www2.census.gov/geo/maps/metroarea/us_wall/Feb2013/necta_0213.pdf on 19 May 2024.
23. **Chandra A.** Code and data documentation for "Mask mandates and COVID-19: a re-analysis of the Boston school mask study." 2024. Accessed at https://github.com/ambarishchandra/MA_maskmandate_reanalysis on 1 August 2024.
24. **Callaway B, Sant'Anna PHC.** Difference-in-differences with multiple time periods. J Econom. 2021;225:200-230. [10.1016/j.jeconom.2020.12.001]
25. **COVID-19 Forecasting Team.** Past SARS-CoV-2 infection protection against re-infection: a systematic review and meta-analysis. Lancet. 2023;401:833-842. [PMID: 36930674] doi:10.1016/S0140-6736(22)02465-5
26. **Schechter-Perkins EM, Doron S, Johnston R, et al.** A test-to-stay modified quarantine program for COVID-19 in schools. Pediatrics. 2022;149:e2021055727. doi:10.1542/peds.2021-055727
27. **Rader B, Gertz A, Iuliano AD, et al.** Use of at-home COVID-19 tests — United States, August 23, 2021–March 12, 2022. MMWR Morb Mortal Wkly Rep. 2022;71:489-494. [PMID: 35358168] doi:10.15585/mmwr.mm7113e1
28. **Branch-Elliman W, Ertem Z, Nelson R, et al.** Impacts of testing, vaccination, and immunity on COVID-19 cases in schools. Research Square. Preprint posted online 16 February 2023. doi:10.21203/rs.3.rs-2587698/v1
29. **U.S. Census Bureau.** American Community Survey–Five-Year Estimates, 2021. Accessed at www.census.gov/data/developers/data-sets/acs-5year.html on 1 August 2024.
30. **Murray EJ, Cowger TL, Hall KT.** Lifting universal school masking — Covid-19 incidence among students and staff. Reply. N Engl J Med. 2023;389:580. [PMID: 37590463] doi:10.1056/NEJMc2215560
31. **Sandlund J, Duriseti R, Ladhani SN, et al.** Child mask mandates for COVID-19: a systematic review. Arch Dis Child. 2024;109:e2. [PMID: 38050026] doi:10.1136/archdischild-2023-326215
32. **Glasziou PP, Michie S, Fretheim A.** Public health measures for covid-19. BMJ. 2021;375:n2729. [PMID: 34789507] doi:10.1136/bmj.n2729
33. **Collins R, Bowman L, Landray M, et al.** The magic of randomization versus the myth of real-world evidence. N Engl J Med. 2020;382:674-678. [PMID: 32053307] doi:10.1056/NEJMsb1901642
34. **Høeg TB, Prasad VK.** An evidence double standard for pharmacological versus non-pharmacological interventions: lessons from the COVID-19 pandemic. Contemp Clin Trials Commun. 2023;33:101108. [PMID: 36938319] doi:10.1016/j.conctc.2023.101108
35. **Hulme WJ, Williamson E, Horne EMF, et al.** Challenges in estimating the effectiveness of COVID-19 vaccination using observational data. Ann Intern Med. 2023;176:685-693. [PMID: 37126810] doi:10.7326/M21-4269
36. **Falk A, Benda A, Falk P, et al.** COVID-19 cases and transmission in 17 K-12 schools — Wood County, Wisconsin, August 31–November 29, 2020. MMWR Morb Mortal Wkly Rep. 2021;70:136-140. [PMID: 33507890] doi:10.15585/mmwr.mm7004e3
37. **Zimmerman KO, Akinboyo IC, Brookhart MA, et al; ABC SCIENCE COLLABORATIVE.** Incidence and secondary transmission of SARS-CoV-2 infections in schools. Pediatrics. 2021;147:e2020048090. [PMID: 33419869] doi:10.1542/peds.2020-048090
38. **Brandal LT, Ofitserova TS, Meijerink H, et al.** Minimal transmission of SARS-CoV-2 from paediatric COVID-19 cases in primary schools, Norway, August to November 2020. Euro Surveill. 2021;26:2002011. doi:10.2807/1560-7917.ES.2020.26.1.2002011
39. **Rotevatn TA, Bergstad Larsen V, Bjordal Johansen T, et al.** Transmission of SARS-CoV-2 in Norwegian schools during academic year 2020-21: population wide, register based cohort study. BMJ Med. 2022;1:e000026. [PMID: 36936575] doi:10.1136/bmjmed-2021-000026
40. **Bundgaard H, Bundgaard JS, 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. A randomized controlled trial. Ann Intern Med. 2021;174:335-343. [PMID: 33205991] doi:10.7326/M20-6817
41. **Howick J, Koletsi D, Ioannidis JPA, et al.** Most healthcare interventions tested in Cochrane Reviews are not effective according to high quality evidence: a systematic review and meta-analysis. J Clin Epidemiol. 2022;148:160-169. [PMID: 35447356] doi:10.1016/j.jclinepi.2022.04.017

Downloaded from https://annals.org by Univ of California San Francisco on 11/08/2024.

Author Contributions: Conception and design: A. Chandra, R. Duriseti, T.B. Høeg, V. Prasad.
Analysis and interpretation of the data: A. Chandra, R. Duriseti, T.B. Høeg, S. Ladhani.
Drafting of the article: A. Chandra, R. Duriseti, T.B Høeg.
Critical revision for important intellectual content: A. Chandra, R. Duriseti, T.B. Høeg, S. Ladhani.
Final approval of the article: A. Chandra, R. Duriseti, T.B. Høeg, S. Ladhani, V. Prasad.
Statistical expertise: A. Chandra, R. Duriseti, T.B. Høeg.
Collection and assembly of data: A. Chandra, R. Duriseti, T.B. Høeg.

Downloaded from https://annals.org by Univ of California San Francisco on 11/08/2024.