COVID-19 Health Policy Impact Evaluation Review

Start of Block: Main form

Q10 Administrative information

Q8 Study DOI

Q3 Reviewer number

Q54 Review type/round

The first round (Primary/independent review round) is for the independent first reviews of every article; the second (Secondary/consensus round) is for the second round of review for each article.

- Primary/independent review round (1)
- Secondary/consensus round (2)

Q50 Screening
Q52 Do you wish to recuse yourself from reviewing this study for any reason (e.g. social or professional relationship with the authors, financial conflict of interest, etc)?

- No, I do not wish to recuse myself. (1)
- Yes, I recuse myself from reviewing this paper. (2)

Skip To: End of Survey If Q52 = Yes, I recuse myself from reviewing this paper.

Q51 Do you believe that this study meets the inclusion criteria for this research?

The inclusion criteria are: The primary topic of the article must be evaluating one or more individual COVID-19 policies on direct COVID-19 outcomes. The primary exposure(s) must be a policy, defined as a government-issued order at any government level to address a directly COVID-19-related outcome (e.g. mask requirements, travel restrictions, etc). COVID-19 outcomes may include cases detected, mortality, number of tests taken, test positivity rates, Rt, etc. This may NOT include indirect impacts of COVID-19 on things such as income, childcare, trust in science, etc. The primary outcome being examined must be a COVID-19-specific outcome, as above. The study must be designed as an impact evaluation study from primary data (i.e. not primarily a predictive or simulation model or meta-analysis). The study must be peer reviewed, and published in a peer-reviewed journal indexed by PubMed. The study must have the title and abstract available via PubMed at the time of the study start date. The study must be written in English.

- Yes (1)
- No (2) ________________________________________________

Skip To: End of Survey If Q51 = No

Q7 Study topic information

Please consult review guidance (available here) for additional guidance on answering these questions.

Q6 Main impact sentence
Copy and paste the sentence from the abstract that best describes the main claim of the study (e.g. "Policy X had a positive impact on outcome Y")
Q9 Main COVID-19 policy type evaluated

Select all that apply. Note: categorization from the Oxford Government Response Tracker

☐ School closing (1)
☐ Workplace closing (2)
☐ Cancel public events (3)
☐ Restrictions on gathering size (4)
☐ Close public transportation (5)
☐ Stay at home requirements (6)
☐ Restrictions on internal movement (7)
☐ Restrictions on international travel (8)
☐ Income support (9)
☐ Debt/contract relief for household (10)
☐ Fiscal measures (11)
☐ Giving international support (12)
☐ Public information campaign (13)
☐ Testing policy (14)
☐ Contact tracing (15)
☐ Emergency investment in healthcare (16)
☐ Investment in COVID-19 vaccines (17)
Q12 Main COVID-19 outcome type evaluated

Select all that apply

☐ COVID-19 cases (1)
☐ COVID-19 test positivity (2)
☐ COVID-19 deaths (3)
☐ COVID-19 hospitalizations (4)
☐ SARS-CoV-2 infections and infection rate (e.g. effective R) (8)
☐ Other (fill in) (9) ________________________________________________

Q13 Method(s) identification

For this section, consider only the data structure as it enters into the main statistical model. In other words, if the original dataset is of individuals at many time points, but the main statistical model uses a regional-level aggregated count of cases, the data as it enters into the main statistical model is a regional aggregate at one time point.

Q14 What is the level of aggregation for the main outcome data?

☐ Individual level (1)
☐ Regional aggregate (e.g. count, mean, etc.) (2)
Q16 How many regional units included in the main statistical model received the policy of interest?

If 2-20, enter the number of regional units analyzed which received the policy of interest.

- One (1) (1)
- Two through twenty (2-20) (2)
- More than twenty (21+) (3)
- Unclear or N/A (4) ________________________________

Q17 How many regional units were included which did NOT receive any form of the policy of interest?

If 2-20, enter the number of regional units analyzed which did not receive the policy of interest.

- Zero (0) (1)
- One (1) (2)
- Two through twenty (2-20) (3)
- More than twenty (21+) (4)
- Unclear or N/A (5) ________________________________

Display This Question:

If Q17 = Zero (0)
Q25 Did different regions receive different intensities of the policy of interest for comparison?

For example, the study might compare places with more intense versions of policy or policies vs. places with less intense versions of policy or policies, rather than just places with and without the policy or policies.

- Yes (regions with more intense policy were compared with regions with less intense policy) (1)
- No (2)
- Unclear or N/A (3)

Q18 For each regional unit, how many time point observations were in the model before the policy was enacted?

- None (0) (1)
- One (1) (2)
- More than one (2+) (3)
- Unclear or N/A (4) ____________________________

Q19 For each regional unit, how many time point observations were in the model after the policy was enacted?

- None (0) (1)
- One (1) (2)
- More than one (2+) (3)
- Unclear or N/A (4) ____________________________
Q20 How would you describe the time intervals between observations?

- Days (1-5 days between observations) (1)
- Weeks (about 5-10 days between observations) (2)
- Multiple weeks (11-25 days between observations) (3)
- Monthly (26 or more days between observations) (4)

Q21 Did the pre-policy period for any region act as a “control” for different region post-policy enactment?

In other words, was there any pre-period in one or more region's being used to control or compare for the trends of any one or more different regions’ post-period?

- No (pre-periods were treated as controls only within-region) (1)
- Yes (pre-periods were treated as controls with other regions) (2)
- Unclear or N/A (3)
Q22 Was any unit assigned the policy or the timing of the policy externally (i.e. as an experiment/trial)?

- No (observational data only) (1)
- Yes (treatment assigned as part of research or evaluation) (2)
- Unclear or N/A (3)

Display This Question:
If Q22 = Yes (treatment assigned as part of research or evaluation)

Q23 Was the assignment randomized?

- Yes (1)
- No (2)

Q27 Based on your answers above and the guidance document, please select the type of study that best resembles the design of the main analysis.

Please note that the design(s) named in the paper may not match with the method described below, nor is this the actual exact design that was used. If you believe that the design used differs from the choices below in a way that makes this choice impossible, please contact the study administrator before selecting "other."

<table>
<thead>
<tr>
<th>Design</th>
<th>Units (e.g., regions of comparison)</th>
<th>Time points measured per unit</th>
<th>Assumed counterfactual.</th>
<th>“If not for the intervention, ___”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>With intervention</td>
<td></td>
<td></td>
<td>With intervention</td>
</tr>
<tr>
<td></td>
<td>Before intervention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>After intervention</td>
<td>Cross-sectional</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At least one</td>
<td>N/A</td>
<td>One time point</td>
<td></td>
</tr>
<tr>
<td></td>
<td>One time point</td>
<td></td>
<td>Outcome in intervention units would have been the same</td>
<td></td>
</tr>
</tbody>
</table>
as the outcome in the non-intervention units.

Pre/post

At least one

None

At least one (typically one)
At least one (typically one)

Outcome would have stayed the same from the pre period to the post period.

Interrupted time-series (ITS)

At least one

None

More than one
At least one (typically several)

Outcome slope and level* would have continued along the same modelled trajectory from the pre-period to the post period.

Difference-in-differences (DiD)

At least one

At least one†

At least one (typically one)
At least one (typically one)

Outcome in intervention units would have changed as much as (or in parallel with) the outcome in the non-intervention units.

Comparative interrupted time series (CITS)

At least one

At least one†

More than one (typically several)
At least one (typically several)

Outcome slope and level* would have changed as much as non-intervention group’s slope and level* changed.

* Assessing both slope and level only applicable if there are multiple data points during the post period † Units without the intervention may be the pre-period of a different unit that eventually receives the intervention.

○ Cross-sectional analysis (1)

○ Non-randomized experiment/trial (2)

○ Randomized controlled trial (3)
Please select one of the following options for your study design:

- Pre/post (4)
- Interrupted time-series (5)
- Difference-in-differences (6)
- Comparative interrupted time-series (7)
- Other (please contact administrator before selecting) (8)

Q49 Design evaluation

Display This Question:
If Q27 = Interrupted time-series
Or Q27 = Difference-in-differences
Or Q27 = Comparative interrupted time-series

Q29 Does the analysis provide graphical representation of the outcome over time?
If not "Yes" please describe (three short sentences max).

- Check for a chart that shows the outcome over time, with the dates of interest, separated by policy/non policy groups if applicable. Outcomes may be aggregated for clarity (e.g. means and CIs at discrete time points).

- Yes (1)
- Mostly yes (2)
- Mostly no (3)
- No (4)
- Unclear (5)
**Display This Question:**

If Q27 = *Interrupted time-series*

Or Q27 = *Difference-in-differences*

Or Q27 = *Comparative interrupted time-series*

Q30 Is there sufficient pre-intervention data to characterize pre-trends in the data?

If not "Yes" please describe (three short sentences max).

- Check the chart(s) to see if there are several time points over a reasonable period of time over which to establish stability and curvature in the pre-trends.

  - Yes (1)
  - Mostly yes (2) ________________________________________________
  - Mostly no (3) ________________________________________________
  - No (4) ________________________________________________
  - Unclear (5) ________________________________________________
Q32 Is the pre-trend stable?

If not "Yes" please describe (three short sentences max).

- Check if there are sufficient data to reasonably determine a stable functional form for the pre-trends, and that they follow a modelable functional form.

  - Yes (1)
  - Mostly yes (2) ________________________________________________
  - Mostly no (3) ________________________________________________
  - No (4) ________________________________________________
  - Unclear (5) ________________________________________________

Display This Question:

If Q27 = Interrupted time-series
Or Q27 = Comparative interrupted time-series

Q31 Is there sufficient post-intervention data to observe post trends in the data?

If not "Yes" please describe (three short sentences max).

- Check the chart(s) to see if there are several time points over a reasonable period of time over which to establish stability and curvature in the post-trends.

  - Yes (1)
  - Mostly yes (2) ________________________________________________
  - Mostly no (3) ________________________________________________
  - No (4) ________________________________________________
  - Unclear (5) ________________________________________________
Q33 Is the functional form of the counterfactual (e.g. linear) well-justified and appropriate?

If not "Yes" please describe (three short sentences max).

-Check whether the authors explain and justify their choice of functional form.
-Check if there is any curvature in the pre-trend.
-Consider the nature of the outcome. Is it sensible to measure the trend of this outcome on the scale and form used? Note: infectious disease dynamics are rarely linear.
-Consider that while pre-trend fit is a necessary condition for an appropriate linear counterfactual model, it is not sufficient. Check if the authors provide justification for the functional form to continue to be of the same functional form (e.g. linear).

- Yes (1)
- Mostly yes (2) ____________________________________________
- Mostly no (3) ____________________________________________
- No (4) _________________________________________________
- Unclear (5) ____________________________________________

Q34 Is the date or time threshold set to the appropriate date or time (e.g. is there lag between the intervention and outcome)?

If not "Yes" please describe (three short sentences max).

-Check whether the authors justify the use of the date threshold relative to the date of the intervention.
-Trace the process between the intervention being put in place to when observable effects in
the outcome might appear over time.
- Consider whether there are anticipation effects (e.g. do people change behaviors before the date when the intervention begins?)
- Consider whether there are lag effects. (e.g. does it take time for behaviors to change, behavior change to impact infections, infections to impact testing, and data to be collected, etc?)
- Check if authors appropriately and directly account for these time effects.

- Yes (1)
- Mostly yes (2)
- Mostly no (3)
- No (4)
- Unclear (5)

Display This Question:
If Q27 = Interrupted time-series
Q36 Is this policy the only uncontrolled or unadjusted-for way in which the outcome could have changed during the measurement period?

If not "Yes" please describe (three short sentences max).

-Consider other policies or interventions which could impact the outcome during this time.
-Consider social behaviors changed which could meaningfully impact the outcome during this time.
-Consider economic conditions changed which could meaningfully impact the outcome during this time.
-Note that the actual concurrent changes do not need to happen during the period of measurement, just their effects.

- [ ] Yes (1)
- [ ] Mostly yes (2) ________________________________________________
- [ ] Mostly no (3) ________________________________________________
- [ ] No (4) ________________________________________________
- [ ] Unclear (5) ________________________________________________

---

**Display This Question:**

If Q27 = Difference-in-differences
Or Q27 = Comparative interrupted time-series

Q53

Is this policy the only uncontrolled or unadjusted-for way in which the outcome could have changed during the measurement period, differently for policy and non-policy regions?

If not "Yes" please describe (three short sentences max).

-Consider any uncontrolled factor which could have influenced the outcome differently in policy and non-policy regions.

-This may include (but is not limited to)
  - Other policies
  - Social behaviors
  - Economic conditions

- Are these factors justified as having negligible impact?
  - If justified, is the argument that these have negligible impact convincing?
-Note that the actual concurrent changes do not need to happen during the period of measurement, just their effects.

-Yes (1)
-Mostly yes (2) ________________________________________________
- Mostly no (3) ________________________________________________
- No (4) ________________________________________________
- Unclear (5) ________________________________________________

Display This Question:
If Q27 = Interrupted time-series
Or Q27 = Difference-in-differences
Or Q27 = Comparative interrupted time-series

Q38
Did authors provide diagnostics or show robustness and/or sensitivity of results to alternative model choices?

If not "Yes" please describe (three short sentences max).

-Yes (1)
- Mostly yes (2) ________________________________________________
- Mostly no (3) ________________________________________________
- No (4) ________________________________________________
- Unclear (5) ________________________________________________
Q39 Given the above, do you believe that the design is appropriate for identifying the policy impact(s) of interest?

This should be taken as independent of what you believe about other studies, and/or the feasibility of other designs.

If not "Yes" please describe (three short sentences max).

☐ Yes (1)

☐ Mostly yes (2) ________________________________________________

☐ Mostly no (3) ________________________________________________

☐ No (4) ________________________________________________

☐ Unclear (5) ________________________________________________

Q55 General and/or additional comments on this paper from consensus discussion. This may include any additional information worth commenting on regarding the paper, difficulties encountered evaluating it, etc.

(three short sentences max)

________________________________________________________________