

# Conceptual lessons for containing an epidemic of questionable COVID-19 counterfactual comparisons

Jonathan Fuller 

The COVID-19 pandemic has provided limitless opportunities to compare pandemic policies across countries and over time. When the aim is to assess the comparative success of these policies, the comparison requires thinking counterfactually about ‘what would have been’ in some unrealised hypothetical (counterfactual) scenario. Whether generating modelling projections,<sup>1</sup> making data-driven comparisons across countries<sup>2</sup> or attributing excess harms,<sup>3</sup> causal inference often rests on counterfactual comparisons, even if those comparisons are only implicit. However, in the pandemic, counterfactual analyses that are overly simplistic, uninformative or outright flawed have been an epidemic in their own right. The examples I explore here are not the worst offenders and my aim is not to criticise them but to use them to illustrate cautionary lessons. By exploring the theory of counterfactuals and common problems with their use, we can learn to do better. Slow conceptual thinking is needed even in times of fast science.

Counterfactuals have played a central role in discussions of causation in philosophy<sup>4</sup> and in the health sciences<sup>5</sup> and social sciences<sup>6</sup> over the past 50 years. According to a framework popular in these disciplines, an intervention causes some outcome if that outcome represents a difference between two hypothetical scenarios in which only the intervention differs. Because the scenarios are mutually incompatible, at least one of them is ‘counterfactual’—that is, contrary to what actually occurs or ‘counter to fact’. Philosophers sometimes think about a counterfactual scenario as an imaginary world that is perfectly identical to the actual world except that the intervention is miraculously altered or manipulated with surgical precision. For instance, if the number of COVID-19 cases would be greater in a possible world that is identical to the real world but in which no pandemic policies were implemented, then we can conclude

that those policies prevented COVID-19 in the actual world.

Scientists and policy-makers cannot make a counterfactual comparison directly because other possible worlds are a fiction (or if they are real then they are inaccessible to us), although they can approximate such a comparison through modelling or using real-world data. A key to doing this well is to first explicitly consider what counterfactual comparison we wish to learn about and then ask what modelling or data would faithfully or usefully represent it. Unfortunately, it is easy to lose sight of the relevance of the available data for the intended counterfactual comparison and of the relevance of the counterfactual comparison for decision-making.

For instance, COVID-19 model predictions have frequently been criticised as inaccurate<sup>7</sup> and no doubt many of them are. However, it is important to distinguish ‘projections’ of what would occur under a hypothetical scenario (which may be counterfactual) from ‘forecasts’ of what will actually occur<sup>8</sup>—a distinction that has not always been marked. Unlike forecasts (such as weather predictions), the accuracy of a counterfactual projection cannot be accurately judged by comparing it to what actually occurred. Schroeder<sup>9</sup> identifies ambiguities in the way that modellers at the Institute for Health Metrics and Evaluation at the University of Washington presented predictions from their epidemic model early on, which sometimes appeared to be projections and sometimes appeared to be forecasts. This kind of ambiguity makes it difficult to evaluate the performance of a model and to know what upshots to draw from its predictions. For instance, while forecasts can help planners anticipate health-care resource usage, projections can help decision-makers choose from among alternative public health policies.<sup>10</sup>

Compartment models like one used by Imperial College London<sup>1</sup> are more clearly ‘projection models’.<sup>8</sup> However, the hypothetical nature of projections allows us to entertain scenarios that realistically would not occur, creating comparisons with questionable relevance for decision-making. In March 2020, Imperial College

modellers claimed that ‘38.7 million lives could be saved’<sup>1</sup> by an aggressive viral-suppression strategy after modelling that scenario (among others) and comparing it to an unmitigated pandemic scenario in which no new actions are taken to contain viral spread. But for evaluating the aggressive suppression strategy, the unmitigated scenario is an unrealistic counterfactual because in that scenario everyone—including governments and the people—behaves as if there were not a pandemic raging. More informative comparisons contrast alternate anticontagion policies or account for the likelihood of evolving anticontagion behaviour even in the absence of aggressive anticontagion policies.

With country-level case data available at a click, many people have made policy comparisons across countries along with inferences regarding the effectiveness of those policies. But comparing one country to another to infer the comparative effectiveness of stricter and laxer (or simply different) anticontagion policies is fraught because it may not faithfully represent a relevant counterfactual comparison.

For example, Bendavid *et al*<sup>2</sup> compared eight countries, including the USA and England, that implemented mandatory stay-at-home orders and business closures with Sweden and South Korea, which did not. To evaluate the effect of these policies on the growth of COVID-19 cases, they subtracted case data in Sweden and South Korea from case data in the other eight countries. In this study, Sweden and South Korea are essentially being used to represent a counterfactual USA or England that does not implement restrictive policies. However, there are important differences between the USA/England and Sweden/South Korea, including social and geographic differences and differences in implementation of other pandemic interventions. Therefore, it seems highly plausible that a cross-country comparison involving the USA or England on one side and Sweden or South Korea on the other fails to accurately represent the outcomes in a ‘USA versus counterfactual USA’ or ‘England versus counterfactual England’ comparison. Other studies (which are by no means infallible) seek to mitigate this problem by making before-and-after comparisons within a country, pooling data from many countries and attempting to adjust for their differences or running sensitivity analyses to test various assumptions.<sup>11 12</sup>

Finally, many have calculated or estimated excess harms in 2020–2021 and beyond such as excess all-cause mortality<sup>13</sup>

History and Philosophy of Science, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Correspondence to Dr Jonathan Fuller, University of Pittsburgh, Pittsburgh, PA 15260, USA; JPF53@pitt.edu

or excess 'deaths of despair'.<sup>14</sup> Excess harms are typically estimated by measuring a stable baseline level of harm (or a stable trend) in recent years and comparing it to the amount of harm measured since the pandemic began or the amount of harm estimated to occur in future years. It is often reasonable to interpret excess harm figures as the increase in harm compared with a counterfactual scenario in which the pandemic never happened. However, it is often more challenging to attribute this increase to a specific factor such as particular policies. Such a harm attribution relies on a different counterfactual comparison between two worlds in which the COVID-19 pandemic is similarly occurring but in which different policies are undertaken. As when measuring beneficial effects, the relevant modelling or data might compare different countries that naturally implemented different policies in 2020–2021 or the same countries before and after the implementation of certain policies.

To illustrate, Niedzwiedz *et al*<sup>3</sup> sought to measure the impact of lockdowns in the UK during 2020 on mental health outcomes through survey results in a longitudinal cohort study. By comparing the prevalence of outcomes such as psychological distress in April 2020 to its prevalence in 2017–2019, they calculated increases or decreases in these outcomes. However, one cannot attribute changes in these outcomes to particular policies from the time trend data alone because, again, in the relevant counterfactual comparison the presence of the pandemic is kept constant and only particular policies are allowed to vary.

Faced with a devastating pandemic rife with examples of countries that followed different paths, regrets about past choices and new decisions to be made, scientists,

policy-makers and the public entertain counterfactual comparisons, comparing what did occur to what would have occurred or what could occur in the future under different scenarios. The ubiquity of models and data available to us makes it possible to provide (more or less reliable) representations of our imagined counterfactual comparisons. But in thinking counterfactually, we must be wary of letting our imagination exceed our data.

**Twitter** Jonathan Fuller @JonathanJFuller

**Acknowledgements** The author thanks Sander Greenland for extensive and thoughtful input on multiple drafts of this manuscript as well as anonymous reviewers.

**Funding** The author has not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Commissioned; externally peer reviewed.

This article is made freely available for use in accordance with BMJ's website terms and conditions for the duration of the covid-19 pandemic or until otherwise determined by BMJ. You may use, download and print the article for any lawful, non-commercial purpose (including text and data mining) provided that all copyright notices and trade marks are retained.

© Author(s) (or their employer(s)) 2021. No commercial re-use. See rights and permissions. Published by BMJ.



**To cite** Fuller J. *J Epidemiol Community Health* Epub ahead of print: [please include Day Month Year]. doi:10.1136/jech-2021-216825

Received 15 April 2021  
Accepted 12 August 2021

*J Epidemiol Community Health* 2021;0:1–2.  
doi:10.1136/jech-2021-216825

**ORCID iD**

Jonathan Fuller <http://orcid.org/0000-0002-7598-9808>

**REFERENCES**

- Walker PGT. The global impact of COVID-19 and strategies for mitigation and suppression, 2020. Available: <https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-26-COVID19-Report-12.pdf> [Accessed 14 Apr 2021].
- Bendavid E, Oh C, Bhattacharya J, *et al*. Assessing mandatory stay-at-home and business closure effects on the spread of COVID-19. *Eur J Clin Invest* 2021;51:e13484.
- Niedzwiedz CL, Green MJ, Benzeval M, *et al*. Mental health and health behaviours before and during the initial phase of the COVID-19 lockdown: longitudinal analyses of the UK household longitudinal study. *J Epidemiol Community Health* 2021;75:224–31.
- Menzies P, Beebe H. Counterfactual theories of causation. In: *The Stanford encyclopedia of philosophy*. Winter 2020 edn, 2020. <https://plato.stanford.edu/archives/win2020/entries/causation-counterfactual/>
- Greenland S. Causal analysis in the health sciences. *J Am Stat Assoc* 2000;95:286–9.
- Morgan SL, Winship C. *Counterfactuals and causal inference: methods and principles for social research*. Cambridge: Cambridge University Press, 2007.
- Ioannidis JPA, Cripps S, Tanner MA. Forecasting for COVID-19 has failed. *Int J Forecast* 2020. doi:10.1016/j.ijforecast.2020.08.004. [Epub ahead of print: 25 08 2020].
- Adams J. What are COVID-19 models modeling? 2020. The Society Pages. Available: <https://thesocietypages.org/specials/what-are-covid-19-models-modeling/> [Accessed 14 Apr 2021].
- Schroeder SA. How to Interpret Covid-19 predictions: reassessing the IHME's model. *Philosophy of Medicine* 2021;2.
- Hernán MA, Hsu J, Healy B. A second chance to get causal inference right: a classification of data science tasks. *CHANCE* 2019;32:42–9.
- Flaxman S, Mishra S, Gandy A, *et al*. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 2020;584:257–61.
- Brauner JM, Mindermann S, Sharma M, *et al*. Inferring the effectiveness of government interventions against COVID-19. *Science* 2021;371:eabd9338.
- Woolf SH, Chapman DA, Sabo RT, *et al*. Excess deaths from COVID-19 and other causes in the US, March 1, 2020, to January 2, 2021. *JAMA* 2021;325:1786.
- Petterson S, Westfall JM, Miller BF. Projected deaths of despair during the coronavirus recession. well being trust, 2020. Available: [https://wellbeingtrust.org/wp-content/uploads/2020/05/WBT\\_Deaths-of-Despair\\_COVID-19-FINAL-FINAL.pdf](https://wellbeingtrust.org/wp-content/uploads/2020/05/WBT_Deaths-of-Despair_COVID-19-FINAL-FINAL.pdf) [Accessed 14 Apr 2021].