

Better Evaluation Using Evidential Pluralism

FROM LEFT TO RIGHT:
MICHAEL WILDE, JO JONES,
ALEX TROFIMOV, JON WILLIAMSON
CENTRE FOR REASONING,
UNIVERSITY OF KENT



“Our central argument is that for some aspects of the pandemic, especially those characterised by a combination of complexity (multiple variables interacting dynamically with a high degree of uncertainty), urgency (decisions needed in days not years) and threat (the consequences of not acting could be catastrophic), mechanistic evidence has been mission-critical and RCTs difficult or impossible. Thousands of lives were likely lost as a result of what was incorrectly claimed to be an “evidence-based” approach – dismissing or downgrading mechanistic evidence, overvaluing findings from poorly designed or irrelevant RCTs, and advocating for inaction where RCT evidence was lacking. The pandemic is an epistemic opportunity for the EBM movement to come to better understand, debate and embrace EBM+.”
(Greenhalgh et al., 2022, p. 253)

The quote left claims (i) that strict adherence to a so-called ‘evidence-based’ approach during the Covid-19 pandemic is likely to have had severe consequences, and (ii) that adequately addressing complex, critical issues requires a more inclusive approach to evidence. In this piece, we offer such an inclusive approach based on the principles of Evidential Pluralism, a recent philosophical account of causal enquiry.

Current Evaluation Approaches

According to current orthodoxy, experimental studies, especially randomized controlled trials (RCTs), and systematic reviews and meta-analyses of RCTs, are the gold standard of evidence. Other kinds of evidence are usually either given less evidential weight or ignored entirely. Beginning in the 1990’s with the evidence-based medicine (EBM) movement ([Sackett et al., 1996](#)), evidence-based practice was a positive development that aimed to provide rigorous, systematic evaluation procedures to ensure the safety and efficacy of medical interventions. The primacy of RCTs is understandable in this context of development. RCTs are suitable for answering relatively simple, focused questions regarding population level associations, and can provide good evidence of the effectiveness of pharmaceutical interventions ([Greenhalgh et al., 2022](#)). However, RCTs are often inadequate for evaluating more complex problems and interventions, including many problems that arise in public health and public policy.

Dissatisfaction with an orthodox approach to evidence-based policy led to the development of theory-based evaluation, which focuses on mechanisms of implementation (Pawson and Tilley, 1997). The primary aim of theory-based evaluation is to assess how an intervention works in a particular context rather than whether an intervention works ([Pawson et al., 2005](#)). Understanding whether an intervention works, however, is often a key concern of policy makers. As a result, theory-based evaluation tends to be seen as a second-best alternative to orthodox evaluation, to be undertaken only when an orthodox evidence-

based evaluation is not feasible. For example, the UK Government's Magenta Book says, 'If none of these [experimental or quasi-experimental] methods seem appropriate, consider Theory-based methods' (HM Treasury, 2020, p. 47).

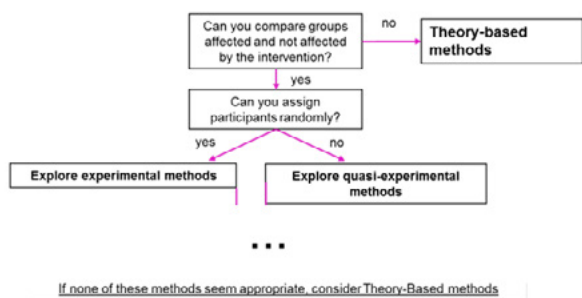


Figure 1: Flowchart on 'Selecting experimental and quasi-experimental methods' from the [UK Government's Magenta Book \(2020, p. 47\)](#).

The time has come for a new approach to evidence-based evaluation. On the one hand, the limitations of orthodox evidence-based practice highlight the need for a more inclusive approach to evidence. On the other hand, although theory-based evaluation can offer important insights into how an intervention works, it does not fully address the question of whether an intervention works. What is needed is an alternative approach to evidence-based evaluation that can tell us what works by systematically incorporating the full range of relevant evidence. Evidential Pluralism offers such an approach.

Evidential Pluralism

Evidential Pluralism is a philosophical account of causal enquiry. Evidential Pluralism includes the following two principles, illustrated in Figure 2:

Object Pluralism: Establishing that A is a cause of B requires establishing (i) a correlation between A and B and (ii) a mechanism connecting A and B that can account for the extent of the identified correlation.

Study Pluralism: Evaluating a causal claim normally requires assessing both association studies and mechanistic studies, where available.

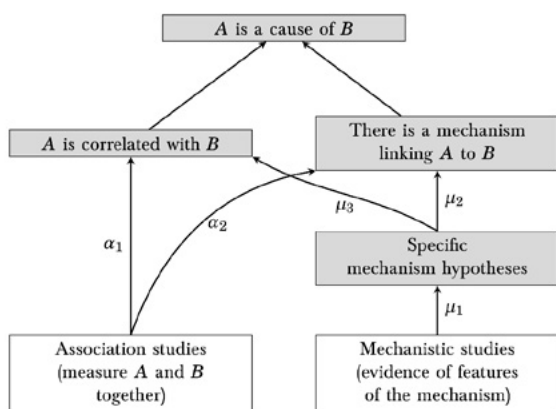


Figure 2: Evidential Pluralism

Motivation for object pluralism comes from the truism that correlation is not causation. A correlation between A and B might be due to A causing B. However, it might be due to any of a large number of alternative explanations, such as confounding, bias, reverse causation or chance. If a correlation between A and B really is due to A causing B, then there must be some mechanism of action by which A produces B. Therefore, establishing causation requires establishing both a correlation and an underlying mechanism that can account for the correlation.

Study pluralism follows from object pluralism. Since establishing causation requires establishing both correlation and mechanism, evaluating a causal claim will normally require assessing both association studies and mechanistic studies. Association studies, according to Evidential Pluralism, test for the presence of a correlation. This is done by repeatedly measuring instances of A and B to determine whether A and B are probabilistically dependent, conditional on potential confounders. Mechanistic studies, according to Evidential Pluralism, include all studies that assess the presence of features of the hypothesised mechanisms linking A and B. Such features might include mediating variables, entities or activities involved in a mechanism, or ways in which these components are organised.

By requiring an assessment of both correlation and mechanisms, Evidential Pluralism provides a framework for integrating diverse kinds of evidence when evaluating whether an intervention works. Assessing correlation requires assessing quantitative evidence from experimental or observational studies. If these studies identify a robust correlation, this can increase confidence that there is some underlying mechanism. However, they will rarely be sufficient on their own to establish that a complex intervention is effective. Evidence of key features of the proposed mechanism of action can increase confidence in effectiveness. On the other hand, if these features are found to be absent, or if key features of counteracting mechanisms are found to be present, this can undermine confidence in effectiveness.

Evidential Pluralism motivates a more inclusive approach to evidence evaluation in medicine, called EBM+ (Parkkinen et al., 2018), policy, called EBP+ (Shan and Williamson, 2023), and law, called EBL+ (Trofimov and Williamson, under review).

Covid-19 Face Masks Case Study

As Greenhalgh et al., 2022 suggest, a narrow focus on experimental studies may be partly to blame for the continued controversy and uncertainty concerning the effectiveness of Covid-19 public health interventions, including public face mask mandates. To illustrate the need for and benefits of an Evidential Pluralism evaluation, we conducted a proof-of-concept evaluation of Covid-19 face mask mandates (Trofimov and Williamson, under review). This evaluation did not include a full systematic literature review and therefore the conclusions should be taken as tentative. Nevertheless, this case study provides a good illustration of the need for and benefits of an Evidential Pluralism evaluation. Here, we provide a summary of that proof-of-concept evaluation.

An Evidential Pluralism evaluation begins by specifying the claims of interest. In this case, they are:

The causal claim: a legal requirement to wear a face mask in public reduces the prevalence of Covid-19 infections and thereby reduces the number of hospitalizations and deaths.

The correlation claim: a legal requirement to wear a face mask in public is negatively correlated with symptomatic Covid-19 infections, conditional on potential confounders.

A plausible mechanism hypothesis: a legal requirement to wear a face mask in public increases the use of face masks which in turn reduces the prevalence of Covid-19 which reduces the prevalence of symptomatic infections and thereby the number of hospitalizations and deaths (Figure 3).

A plausible counteracting mechanism hypothesis: a legal requirement to wear a face mask in public will decrease compliance with other public health interventions, such as social distancing. This, in turn, would result in an increase in the number of symptomatic infections compared to the number that would have occurred if the legal requirement to wear a face mask in public had not been implemented (Figure 3).

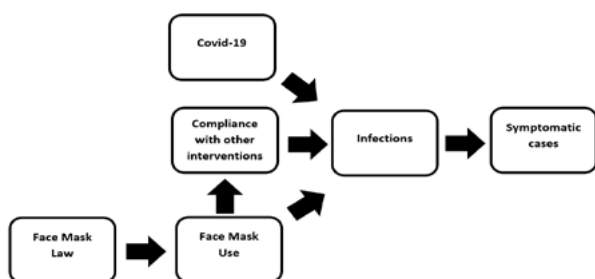


Figure 3: Mechanism Hypothesis.

The next stage is to search for and assess association studies, to enable an evaluation of the correlation claim. The search found a range of association studies that, taken together, detect a robust correlation across contexts. On the basis of that evidence, Trofimov and Williamson conclude that the correlation claim is established and that confidence is increased in the existence of a mechanism that accounts for this correlation. Although each of the association studies controlled for several confounders, some residual confounding may persist. The association studies therefore do not suffice to fully establish causation and a mechanistic evaluation was deemed necessary because it could further increase confidence in the causal claim.

A literature search revealed that each stage of the mechanism hypothesis is supported by a range of studies and that there is evidence against the hypothesised counteracting mechanism. On the strength of the evidence considered, Trofimov and Williamson conclude that the mechanism hypothesis is confirmed but not fully established, and that the counteracting mechanism hypothesis is arguably false.

Overall, the evaluation found strong evidence of both correlation and mechanisms. Association studies on their own establish the correlation claim and the mechanistic studies reinforced this conclusion. Neither association studies on their own nor mechanistic studies on their own fully establish the existence of a mechanism that can account for the correlation.

When association studies and mechanistic studies are considered together, however, confidence in the existence of such a mechanism is increased to such an extent that it is unlikely that further evidence would undermine it. Therefore, on the basis of the mutually supporting evidence of correlation and mechanism, the mechanistic claim was deemed to be established. Since both the correlation claim and the mechanistic claim are established, Evidential Pluralism implies that the causal claim is established.

By taking account of a broad range of evidence, Trofimov and Williamson were able to reach a robust, positive conclusion regarding the effectiveness of Covid-19 face mask mandates. Given the continued controversy and uncertainty surrounding Covid-19 face mask mandates, this is significant. For a more detailed explanation and further examples, please see our introductory guide (Jones, Trofimov, Wilde and Williamson, 2024).

Conclusion

As the pandemic has highlighted, there is a pressing need for a more inclusive approach to evidence-based evaluation, especially for complex issues. While this is becoming more widely recognised, official guidance and practice does not offer a framework for systematically integrating different kinds of evidence to yield better informed evaluations of effectiveness. Evidential Pluralism offers such a framework and thereby promises to improve evidence-based evaluation.

References

- Greenhalgh, T., Fisman, D., Cane, D., J., Oliver, M. and Macintyre, C. R., (2022), 'Adapt or Die: How the Pandemic made the shift from EBM to EBM+ more Urgent', *BMJ Evidence-Based Medicine*, 27: 253-260.
- HM Treasury (2011), *Magenta Book: Central Government Guidance on Evaluation*, HM Treasury and Evaluation Task Force, London, 2020 edition. Available at: [The Magenta Book - GOV.UK \(www.gov.uk\)](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/214242/magenta-book-2011.pdf)
- Jones, J., Trofimov, A., Wilde, M. and Williamson, J., (2024) 'Integrating Diverse Evidence Using Evidential Pluralism: Introductory Guide'. Available at: [IDE-EP.pdf \(jonwilliamson.uk\)](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/123456/IDE-EP.pdf)
- Parkkinen, V-P, Wallmann, C., Wilde, M., Clarke, B., Illari, P., Kelly, M. P., Norell, C., Russo, F., Shaw, B. and Williamson, J., (2018), *Evaluating Evidence of Mechanisms in Medicine: Principles and Procedures*, Switzerland: Springer.
- Pawson, R., Greenhalgh, T., Harvey, G and Walshe, K., (2005), 'Realist Review – A New method of Systematic Review Designed for Complex Policy Interventions', *Journal of Health Services Research and Policy*, 10: 21-34.
- Pawson, R. and Tilley, N., (1997), *Realistic Evaluation*, London: Sage.
- Sackett, D. L., Rosenberg, W. M., C., Gray, J. A. M., Haynes, R. B. and Richardson, W. S., (1996), 'Evidence Based Medicine: What It Is and What It Isn't', *BMJ*, 312: 71.
- Shan, Y. and Williamson, J., (2023), *Evidential Pluralism is the Social Sciences*, Abingdon: Routledge.
- Trofimov, A. and Williamson, J., (under review), 'Applying Evidential Pluralism to Evidence-Based Law: EBL+'.