

**Dialing back ‘impact’ claims: researchers should not be compelled to make policy claims based on single studies**

David Bann<sup>1\*</sup>, Emilie Courtin<sup>2</sup>, Neil M Davies<sup>3-4</sup>, Liam Wright<sup>1</sup>

<sup>1</sup>Centre for Longitudinal Studies, Social Research Institute, UCL, London, UK

<sup>2</sup>Department of Health Policy, LSE, London, UK

<sup>3</sup>Division of Psychiatry and Department of Statistical Science, University College London, UK

<sup>4</sup>K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Norway

Word count: 1942

\*Corresponding author. David Bann PhD, Centre for Longitudinal Studies, University College London, London, UK. E-mail: david.bann@ucl.ac.uk

Key words: Policy Implications, Research Impact, Evidence-based Policy, Academic Incentives, Epidemiology.

Researchers are increasingly expected to draw policy implications from their research, yet this can be distracting or misleading when describing single studies. Rather than helping ensure that research benefits society, it may distort research and evidence-based policy. This is propelled by incentives such as career structures increasingly favoring evidence of ‘impact’ and a need to appeal to competitive publication and funding decisions. We discuss this issue, use an example from the health inequality literature to highlight the complications of drawing policy conclusions, and consider options to improve both research and evidence-based policy making.

## **Background**

The incentive structures in science are problematic.<sup>1 2</sup> Fraud cases are numerous, results are often overhyped, and major research papers across many fields are not reproducible.<sup>1 2</sup> In some ways the problems appear to be getting worse. Recent work has found the use of ‘hype’ or positive language has increased across time in scientific articles<sup>3 4</sup> and funded grant applications.<sup>5</sup>

There are also increasing financial and career incentives to produce research judged to lead to broader societal (policy or political) impact. Take, for example, the UK's impact agenda—substantial financial returns are tied to ‘impact’ (25% of University core funding<sup>6</sup>), defined as the “demonstrable contribution that excellent research makes to society and the economy”.<sup>7</sup> This agenda ostensibly rewards the broader contribution of research unlike narrower indicators of research productivity (e.g., citation metrics), yet there are increasing concerns about its downsides.<sup>8</sup> Unsurprisingly given the large incentive, hype is common in impact submissions,<sup>9</sup> and universities spend significant sums on impact generation (at the expense of research).<sup>9 10</sup> There is increasing awareness of the importance of impact for scientific career progression, including awards provided by research funders. Editors of journals operate in an increasingly competitive landscape, and authors must make their articles appealing to publish in ‘top’ outlets.

Papers or grant proposals which purport to advance policy may be preferred to those which merely 'add' to the scientific literature;<sup>11</sup> outputs from such research may garner more media and policy interest, and ultimately might lead to 'demonstrable' impact (e.g., being cited in a policy report, attracting media attention, or being a claimable cause of policy change).

We recognise the importance of policy-relevance of research (the questions we ask and try to answer) but we are concerned that misaligned incentives are leading to researchers increasingly drawing—or at least being compelled to draw—unfounded policy conclusions based on single studies. This is despite an increasing understanding of causality (the 'causal revolution')<sup>12 13</sup> and previous calls to promote evidence synthesis and triangulation rather than individual studies.<sup>14 15</sup> Our concerns are far from new. For example, Rothman and Poole in 1985 warned that "the conduct of science to achieve political ends will corrupt both endeavors".<sup>16</sup> As of 1993, *Epidemiology* explicitly requested that policy implications are contained in commentaries and not in original research articles; illustrating the wider draw of 'impact'; however, more recently established journals which encourage the opposite have considerably higher impact factors (e.g., *Lancet Public Health*—its impact factor is reportedly almost 10 times that of *Epidemiology*).

Researchers are incentivized to follow an alluring structure: describe results of a single study, briefly outline its limitations, and then go on to make broad policy implications. Such pronouncements often appear in the Discussion (Implications subsection) and sometimes in the main abstract. This is to be expected: we are influenced by upstream processes and incentive structures. We have not cited examples, as it seems unfair to pick out individual authors; we are all exposed to this force, as an audit of our own previous publications would demonstrate.

In the context of work which tackles sensitive topics—such as childhood obesity, health inequalities, youth mental health, or social mobility—the lack of political will to address key policy issues is an additional driver of this trend. It leads to frustration among authors, reviewers, editors, and grant committees with merely *describing* associations, encouraging more ambitious policy or political conclusions. However, when articles make unfounded policy conclusions based on single studies, they risk undermining trust in the discipline and science as a whole.

Below, we describe how even strong associations between socioeconomic factors and health have a surprising number of complications which make drawing straightforward policy conclusions challenging. This leads us to suggest how we, as authors, should report on, and as reviewers, should provide feedback on, policy claims made in future.

### **Policy implications of associations between socioeconomic factors and health**

The study of how social disadvantage associates with worse health is an important topic across many disciplines—notably (social) epidemiology and public health, medicine, demography, and economics. Work in this field is often motivated by a noble desire to raise awareness of and ameliorate a particularly pervasive form of injustice: those with fewer resources in society are at higher risk of ill health and premature death. These links have been repeatedly documented for decades,<sup>17 18</sup> motivating policy nationally and transnationally.

A substantial share of this research is necessarily descriptive in nature, since randomized controlled trials, as used in medicine, are often impractical or unethical. However, researchers reporting descriptive work in this area often include (or are compelled to include) broad policy implications. This is in contrast to the evidence-based medicine approach, in which clinical recommendations result from systematic reviews and meta-analyses; single studies (e.g., RCTs

of drugs) often focus solely on the study description rather than engaging in broad policy discussion of what the results might mean for future policy. The policy implications of links between social disadvantage and health are complex—they depend on causal processes we have imperfect evidence about and which are likely context-specific. They also require difficult policy considerations more broadly. For example, consider the following illustrative inference of only 42 words, as is often necessary given tight word limits:

“We reported a strong social gradient in health; those with lowest education attainment had a 3-fold higher risk of cardiovascular disease compared with those of highest education. Policies to improve education attainment could reduce health inequalities and benefit overall population health in future.”

Complications implicit in this kind of statement but often not discussed are plentiful:

1. There is a causal link between education and health—a suggestion supported in traditional observational studies but with mixed findings in the (quasi-)experimental literature.<sup>19</sup>
2. The beneficial effects of education will be translated when intervened on in the future (i.e., the consistency assumption<sup>20</sup> is met and effects generalize across different contexts). Policy initiatives to increase the length of education may have effects that differ by context (i.e., time and place). In one example in which longer schooling was not associated with improved educational outcomes, greater education was linked with worse health outcomes for those from lower social classes.<sup>21</sup>
3. The exposure of interest (education) can be translated into a policy intervention: it is possible to intervene on the educational gradient. It remains unclear which dimension of education actually matters for health. As noted above, changes in the duration (quantity)

of schooling have produced mixed effects on health and the literature on changes in the quality of education (e.g. class sizes, curriculum or tracking by ability) is in its infancy.<sup>22</sup>

4. Policies to improve education attainment will not have other unexpected negative consequences to public health (e.g., mental health) or on other societal (e.g., economic) outcomes.
5. The way the effect of interest is presented is key. In our example, only the relative risk was presented; knowledge of the absolute risk of the outcome in each group is typically needed to inform policy decisions.

Further, a broader discussion of the different policy options is not included. Any policy decision necessarily has opportunity costs and trade-offs. For example, the choice between investing more in education, healthcare or social care. Assessing likely costs and benefits of policies are complex analyses that must be done carefully. This key issue which faces those making policy decisions is not explicitly addressed in the above statement. This is true even when decisions have to be made quickly on limited evidence (as happened during COVID, for example).

A more cautious short-form alternative could be provided, which may well be considered uninteresting by those seeking impact from individual studies. For example: "Assuming links are replicable and causal, translate when intervened on, and do not have other negative consequences, increasing educational attainment could be one policy option to reduce health inequalities and improve overall population health. Further research is required to test these assumptions and to conduct cost-benefit analyses of the multiple policy options available."

Note, these issues are also relevant for systematic reviews: descriptive studies do not become any less descriptive simply because they are combined. They also largely apply to quasi-experimental or experimental studies: even with less potential for confounding, policy conclusions aren't always clear-cut. Results may not transfer between settings, there may be

unaccounted-for costs or benefits, and the overall gains from the policy may not surpass its opportunity costs.

Standard research articles in epidemiology/public health are concise, with word counts of 3000-3500 words. Researchers are often in the unfortunate position of being expected to draw policy implications from their research at extreme brevity. In the internet era, concerns regarding word limits will surely be less relevant; if authors want to include policy discussion, editors should consider letting them have more space to do so comprehensively. If they are included, they should be written carefully, and humbly,<sup>11</sup> drawing on evidence from multiple sources. If authors decide not to include such discussion and focus instead on other important aspects of their work, this should not be penalized by editors, reviewers, funding panels, or academic career structures. This applies to authors of both original research articles and review articles.

Researchers and reviewers often lack training or expertise in policy analysis or cost-benefit analysis; future students could be trained in such work, so that researchers working on all types of study designs (1) are aware of the assumptions required to infer causality from their work, (2) understand how their research could contribute to evidence based policy, and (3) can identify policies which can be appropriately analysed with available data.

Researchers have increasingly constrained bandwidths, and it is debatable whether a brief discussion is helpful for many (or indeed most) single studies; if policy discussions are not included, other valuable aspects can be expanded (e.g., more detailed introductions, methods, results or discussion sections; or triangulating their findings across multiple datasets), or new resources provided (e.g., providing analytical code which only ~2% of researchers do<sup>23</sup>). Policy discussion could instead be contained within separate pieces of work, or it could be accompanied separately alongside original research articles; the latter is the approach taken in social science genetics research, with policy-relevant Frequently Asked Questions (FAQs)

appended separately to avoid misinterpretation of their results.<sup>24</sup> This may be optimal given the challenging and specialized nature of contemporary research; it may yield more robust, reproducible science and improve evidence-based policy.

Policy recommendations should generally be based on literatures, not individual papers, and on the myriad of considerations noted above that policy decisions require. Researchers are often motivated by a noble desire to improve society and influence policy. And yet ironically, policy discussion articles are typically less readily publishable or valued than original research articles. If academic research aims to affect policy, separate detailed policy discussion pieces or reviews should be valued by editors, funders, and academic career structures. Policy discussion pieces should be valued as separate pieces of work where they can consider causality, generalizability, and policy trade-offs in sufficient depth. In contrast, researchers should be *less* compelled to make policy claims based on single studies. All types of study, whether observational or experimental, have value in this context, even if individually they cannot ordinarily lead to broader impact.

### **Author contributions**

DB wrote the first draft; all author contributed to discussions, revisions and approved the final version.

### **Funding**

DB and LW are supported by the Economic and Social Research Council (grant number ES/M001660/1); DB, LW and NMD by the Medical Research Council (MR/V002147/1). EC is supported by a UKRI Pioneer grant (EP/Y010345/1). NMD is supported by a Norwegian Research Council Grant number 295989.



## Conflict of Interest

None declared.

## References

1. Ritchie S. Science fictions: Exposing fraud, bias, negligence and hype in science: Random House 2020.
2. Goldacre B, Farley R. Bad science: Fourth Estate London 2009.
3. Hyland K, Jiang FK. 'Our striking results demonstrate...': Persuasion and the growth of academic hype. *Journal of Pragmatics* 2021;182:189-202.
4. Vinkers CH, Tjeldink JK, Otte WM. Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: retrospective analysis. *BMJ* 2015;351
5. Millar N, Batalo B, Budgett B. Trends in the use of promotional language (hype) in abstracts of successful national institutes of health grant applications, 1985-2020. *JAMA Network Open* 2022;5(8):e2228676-e76.
6. Martin BR. The Research Excellence Framework and the 'impact agenda': are we creating a Frankenstein monster? *Research evaluation* 2011;20(3):247-54.
7. UKRI. Defining impact 2023 [
8. Smith KE, Bandola-Gill J, Meer N, et al. The impact agenda: Controversies, consequences and challenges: Policy Press 2020.
9. Hyland K, Jiang F. Hying the REF: promotional elements in impact submissions. *Higher Education* 2023:1-18.
10. Durrant H, MacKillop E. University policy engagement bodies in the UK and the variable meanings of and approaches to impact. *Research Evaluation* 2022;31(3):372-84.
11. Hoekstra R, Vazire S. Aspiring to greater intellectual humility in science. *Nature Human Behaviour* 2021;5(12):1602-07.
12. Pearl J, Mackenzie D. The book of why: the new science of cause and effect: Basic Books 2018.
13. Angrist JD, Pischke J-S. The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *J Econ Perspect* 2010;24(2):3-30.
14. Wilson P, Petticrew M. Why promote the findings of single research studies? *BMJ* 2008;336(7646):722-22.
15. Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005;2(8):e124.
16. Rothman KJ, Poole C. Science and policy making: American Public Health Association, 1985:340-41.
17. Mackenbach JP. Health inequalities: Persistence and change in modern welfare states: Oxford University Press, USA 2019.
18. Department of Health. The Black Report. Inequalities in Health. Report of a Research Working Group. London, UK, 1980.
19. Hamad R, Elser H, Tran DC, et al. How and why studies disagree about the effects of education on health: A systematic review and meta-analysis of studies of compulsory schooling laws. *Soc Sci Med* 2018;212:168-78.
20. Rehkopf DH, Glymour MM, Osypuk TL. The consistency assumption for causal inference in social epidemiology: when a rose is not a rose. *Current epidemiology reports* 2016;3(1):63-71.
21. Courtin E, Nafilyan V, Avendano M, et al. Longer schooling but not better off? A quasi-experimental study of the effect of compulsory schooling on biomarkers in France. *Soc Sci Med* 2019;220:379-86.

22. Glymour MM, Manly JJ. Compulsory Schooling Laws as quasi-experiments for the health effects of education: Reconsidering mechanisms to understand inconsistent results. *Social science & medicine (1982)* 2018;214:67-69.
23. Hamilton DG, Hong K, Fraser H, et al. Prevalence and predictors of data and code sharing in the medical and health sciences: systematic review with meta-analysis of individual participant data. *BMJ* 2023;382:e075767. doi: 10.1136/bmj-2023-075767
24. Martschenko DO, Domingue BW, Matthews LJ, et al. FoGS provides a public FAQ repository for social and behavioral genomic discoveries. *Nat Genet* 2021;53(9):1272-74.
25. Cornfield J, Haenszel W, Hammond EC, et al. Smoking and lung cancer: recent evidence and a discussion of some questions. *J Natl Cancer Inst* 1959;22(1):173-203.