Statistical Science
2020, Vol. 35, No. 3, 476–478
https://doi.org/10.1214/20-STS776
Main article: https://doi.org/10.1214/19-STS728

A New Template for Empirical Studies: From positivity to Positivity

Rhian Daniel

1. GENERAL COMMENTS

The article by VanderWeele et al. [3] is a *tour de force* of innovation, education and pragmatism, and is a must-read for all students and researchers in statistical science and associated disciplines. As well as suggesting a new outcome-wide approach to the analysis of empirical studies, the article gathers together numerous other new or recent methodological and practical advances that are useful even outside the proposed framework. These include the modified disjunctive cause criterion for confounder selection, the E-value for sensitivity analysis, valuable new insights on well-known correction methods for multiple testing and a novel metric for the expected number of false positive findings. In addition, the paper comprehensively summarises the vast literature on confounder-adjusted analyses in a succinct, accessible and educational manner oozing with practical advice, even on how to report results in space-limited journals. As if consciously practising what they preach, the authors include in a single paper an exploration of almost all the associated issues, caveats, extensions, modifications and comparisons: a bells-and-whistles-wide methodological contribution that other authors may have split into a dozen papers or more.

When it comes to issues such as the so-called replication/reproducibility crisis and formal causal aspects of the analysis of observational studies, the awareness of problems and potential pitfalls are of course essential in engendering appropriate caution and humility. However, a *can't do* attitude ("$p < 0.05$ doth not a finding make", "correlation is not causation") is less likely to improve matters than a clear, concrete and implementable alternative approach, such as this one by VanderWeele et al.

Their central suggestion in a nutshell is that researchers concerned with a particular exposure should study (and report) its effect, not on a single outcome of interest to them, but rather on as wide a range of outcomes as is feasible. They discuss the advantages of doing so, many of which relate to research efficiency, reproducibility and lessening publication bias (by ensuring that more null results are published). An obvious but compelling advantage arises when an exposure (e.g., HRT) has a harmful effect on one outcome (e.g., cancer) and a protective effect on another (e.g., heart disease). Such results reported separately contribute to the adage that "today's poison is tomorrow's wonder drug". Many less obvious advantages, for example that some outcomes may plausibly serve as negative controls for others, are also compelling.

The article is (cautiously) positive about the extent to which well-conducted observational studies can offer evidence on cause–effect relationships, with many enlightening and nuanced discussions on the likely magnitude of biases arising from various sources in different contexts. In the remainder of this commentary, I will first mention a few additional minor notes of caution that came to mind whilst attempting to read the article under a commentator's hat. One of these concerns potential violations of the *positivity* assumption (the large-P Positivity in the title), which is discussed in more detail in Section 3. My overwhelming feeling towards the paper, however, is positive (the small-p positivity in the title).

2. A FEW MINOR CAUTIONARY NOTES

In several parts of the article, the authors describe situations in which an analyst will be faced with two suboptimal options. For example, in Section 2.5, when studying the effect of physical activity on cardiovascular disease, BMI is plausibly both a confounder and a mediator. If repeated measurements of BMI are available, one option is to adjust for BMI at a wave previous to the exposure measurement, risking residual confounding, and another is to adjust for a more recent measurement, risking adjusting for a partial mediator of the effect of interest. The authors suggest doing both in a sensitivity analysis. In this example, failure to fully adjust for the confounding through BMI will likely lead to *over*estimating the beneficial effect of physical activity, whereas adjustment for a measurement of BMI on the causal pathway from physical activity to cardiovascular disease will likely lead to *under*estimating the beneficial effect. The two estimates could thus plausibly be viewed as bounds, with the true effect lying somewhere between the two. It is worth mentioning as an additional caution, however, that in many other settings both analyses could be biased in the *same* direction. This will tend to happen whenever the covariate's effect on the exposure is in the opposite direction from the exposure's effect on the covariate (e.g., high

Rhian Daniel is a Reader in Medical Statistics, Division of Population Medicine, Cardiff University, Cardiff, United Kingdom (e-mail: danielr8@cardiff.ac.uk).