

Andrew Gelman’s contribution to the discussion of “Randomisation in clinical trials: Is a pragmatic compromise between randomisation- and model-based inference possible?” by Senn

Andrew Gelman, Professor, Department of Statistics and Department of Political Science,
Columbia University, New York, ag389@columbia.edu

I agree with Senn (2026) that randomization and model-based inference should be viewed as complements, not competitors. Modeling can make randomization-based inference more efficient, and, as explained by Rubin (1978), randomization can make model-based inference more robust.

One reason for taking randomization as a starting point is that, when the experimental units have no distinguishing characteristics, treatment assignment is necessarily exchangeable and is thus equivalent to some form of randomization. In real life, people always have distinguishing characteristics (even if personal information is redacted, the researcher can still have access to distinguishing meta-data such as the date and time of entry into the study), and randomization can require some effort. The point here is that departures from randomization are directly tied to observable differences between units.

I also want to share two examples to demonstrate Senn’s point that, even in randomized designs, much can be gained by adjusting for pre-treatment differences between treatment and control groups. Gelman, Carlin, and Nallamothu (2019) discuss a randomized clinical trial of medical procedure that was analyzed using gain scores—the change in treadmill time for patients before and after the treatment period—yielding an estimate of 16.6 seconds with a standard error of 12.7. It is better to include the pre-treatment measure as a regression predictor—the resulting estimate is unbiased and with lower variance—and in this case the estimate is 21.3 with standard error 12.5, which we think is a more accurate summary of the evidence from the study. For a second example, Gelman (2022) discusses a regression discontinuity analyses comparing historical political candidates where the outcome was remaining years of life. This was an observational study but it suffered from the same problem of not adjusting for a relevant pre-treatment predictor—but even more so, as the available but unused information was politician’s age at the time of the treatment. Unsurprisingly, including age induces a huge improvement in the regression—its estimated coefficient is around -0.9 and R-squared increases from 5% to 31%. One of my two examples is a randomized experiment and the other is an observational study, but in both cases the original researchers relied so heavily on design as to ignore obvious information that could easily be included in a model. From the other direction, all the modeling in the world can’t fix bias from unmeasured confounders. As Senn says, design and analysis and analysis are complementary.

References

Gelman, A. (2022). Criticism as asynchronous collaboration: An example from social science research. *Stat* 11, e464.

Gelman, A., Carlin, J., and Nallamothu, B. (2019). Objective Randomised Blinded Investigation With Optimal Medical Therapy of Angioplasty in Stable Angina (ORBITA) and coronary stents: A case study in the analysis and reporting of clinical trials. *American Heart Journal* 214, 54-59.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics* 6, 34-58.

Senn, S. (2026). Randomisation in clinical trials: Is a pragmatic compromise between randomisation- and model-based inference possible? *Journal of the Royal Statistical Society B*.