Difference in difference, controlled interrupted time series and synthetic controls

James Lopez Bernal[1,2], Steven Cummins[1], Antonio Gasparrini[1,3]

1. Department of Public Health Environments and Society,
2. Immunisations and Countermeasures Division, Public Health England
3. Centre for Statistical Methodology

We thank Benmarhnia and Rudolph [REF] for their critical appraisal on our recent article on the use of controls in interrupted time series (ITS) studies.[1] This offers the opportunity to clarify some important issues related to ITS and controlled ITS (CITS) designs and their comparison with other methods applied for public health evaluation. In particular, we argue that Benmarhnia and Rudolph based their assessment on three incorrect premises: that ITS without control is not a valid design for assessing causal relationships; that CITS is just another name for the difference-in-difference (DID) design that they advocate; that the synthetic control methodology represents an alternative to CITS.

A DID design typically refers to a controlled before and after study in which the outcome is measured at a single baseline (pre-intervention) time point and a single post-intervention time point, or where pre- and post-intervention means are compared but where 'time' is not incorporated into the model.[2-4] The counterfactual is estimated based on the control group alone and assumes that trends are parallel in the two groups. While approaches can be used to ensure that the two groups are as similar as possible, such as using synthetic controls, the assumption that trends are parallel is not verifiable using this design therefore it is highly susceptible to confounding due to between group differences.

ITS, conversely, uses multiple consecutive pre- and post-intervention observations in a single population and incorporates time. The counterfactual is estimated by extrapolation of the pre-intervention trend and assumes that in the absence of an intervention the trend would remain constant. Because the observations are undertaken in the same population, between group differences are do not present a problem, and the strict temporal structure allows fine control for underlying trends and measured time-varying confounders. However, other events occurring around the time of the intervention, can be a source of confounding.[5-7]

Both ITS and DID are generally regarded as intermediate designs in the hierarchy of quasi-experimental designs.[4] However, CITS combines the ITS design with one or more control series, allowing both within and between group comparisons and strengthening the control for potential confounders. As such, it provides a more flexible and structured inferential framework, and it is regarded as a more powerful design that DID.[4,7,8] Notably, as an extension of DID, it allows the assumption of parallel trends to be verified and for differences in trend between two groups to be

adjusted for. As an extension to ITS, it allows time varying confounders, including contemporaneous events, that affect both groups to be controlled for.[1]

Synthetic controls are not an alternative approach to CITS, rather, as we describe, they can used to identify a suitable control series for use in an CITS, and they are thus complementary approaches.[1] The use of synthetic controls in CITS studies has been described in detail elsewhere.[9] Nevertheless, synthetic controls rely on the availability of multiple suitable controls with various measures on several characteristics, a scenario which is often not available in practice.

There are scenarios where no suitable unaffected control group exists, for example, evaluation of national or international policies or events, such as the impact of smoking cessation legislation or the financial crisis, therefore a CITS design would not be possible.[10,11] Benmarhnia and Rudolph suggest that under such circumstances uncontrolled ITS is not a suitable alternative and that the intervention should not be evaluated if an "approach with identification assumptions that are more closely aligned with the data" is not possible". We disagree. Interrupted time series is the most powerful study design available where no control group exists.[4,5] Of course, the evidence emerging from an ITS study has to interpreted in a broader context, taking into account biological plausibility, magnitude of effect and consistency across settings.[12] Further, the inability to control for possible contemporaneous events, should be explicitly acknowledged. Nevertheless, this is not an excuse not to evaluate using the best available approach. Lack of evaluation of such interventions would simply perpetuate the evaluative bias that exists with complex interventions that are challenging to study.[2]

References:

1.Lopez Bernal JA, et al. The Use of Controls in Interrupted Time Series Studies of Public Health Interventions. Int J Epidemiol. 2018;Epub ahead of print.
2.Craig P, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. JECH. 2012;66:1182-6.
3.Rockers PC, et al. Inclusion of quasi-experimental studies in systematic reviews of health systems research. Health Policy. 2015;119:511-21.
4.Soumerai SB, et al. How Do You Know Which Health Care Effectiveness Research You Can Trust? A Guide to Study Design for the Perplexed. Preventing Chronic Disease. 2015;12.
5.Wagner AK, et al. Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther. 2002;27:299-309.
6.Lopez Bernal J, et al. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol. 2016;46:348-55.
7.Shadish WR, et al. Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin; 2002.
8.Lopez Bernal JA, et al. The use of quasi-experimental designs for vaccine evaluation. Clin Infect Dis. 2018.
9.Linden A. Combining synthetic controls and interrupted time series analysis to improve causal inference in program evaluation. Journal of evaluation in clinical practice. 2018.
10.Barone-Adesi F, et al. Effects of Italian smoking regulation on rates of hospital admission for acute coronary events: a country-wide study. PLoS One. 2011;6:e17419.
11.Laliotis I, et al. Total and cause-specific mortality before and after the onset of the Greek economic crisis: an interrupted time-series analysis. The Lancet Public health. 2016;1:e56-e65.

12.Hill AB. The Environment and Disease: Association or Causation? Proc R Soc Med. 1965;58:295-300.