Letter to the Editor

Difference in difference, controlled interrupted time series and synthetic controls

James Lopez Bernal,1* Steven Cummins1 and Antonio Gasparrini1,2

1Department of Public Health Environments and Society, London School of Hygiene and Tropical Medicine, London, UK and 2Centre for Statistical Methodology, London School of Hygiene and Tropical Medicine, London, UK

*Corresponding author. Department of Public Health Environments and Society, London School of Hygiene and Tropical Medicine, London, UK. E-mail: james.lopez-bernal@lshtm.ac.uk

We thank Benmarhnia and Rudolph1 for their critical appraisal of our recent article on the use of controls in interrupted time series (ITS) studies.2 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; and 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.3–5 The counterfactual is estimated based on the control group alone and assumes 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 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.

Both ITS and DID are generally regarded as intermediate designs in the hierarchy of quasi-experimental designs.5 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 than DID.5,8,9 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.2

Synthetic controls are not an alternative approach to CITS; rather, as we describe, they can be used to identify a suitable control series for use in a CITS, and they are thus complementary approaches.2 The use of synthetic controls in CITS studies has been described in detail elsewhere.10 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.11,12 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. ITS is the most powerful study design available where no control group exists.5,6 Of course, the evidence emerging from an ITS study has to be interpreted in a broader context, taking into account biological plausibility, magnitude of effect and consistency across settings.13 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.3

References