Methodological Progress Note: Interrupted Time Series

Sanjay Mahant, MD1,2,3*, Matthew Hall, PhD4

1Division of Pediatric Medicine, Department of Pediatrics, University of Toronto, Toronto, Canada; 2Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada; 3Child Health Evaluative Sciences, Research Institute, Hospital for Sick Children, Toronto, Canada; 4Research and Statistics, Children’s Hospital Association, Lenexa, Kansas.

Hospital medicine research often asks the question whether an intervention, such as a policy or guideline, has improved quality of care and/or whether there were any unintended consequences. Alternatively, investigators may be interested in understanding the impact of an event, such as a natural disaster or a pandemic, on hospital care. The study design that provides the best estimate of the causal effect of the intervention is the randomized controlled trial (RCT). The goal of randomization, which can be implemented at the patient or cluster level (eg, hospitals), is attaining a balance of the known and unknown confounders between study groups.

However, an RCT may not be feasible for several reasons: complexity, insufficient setup time or funding, ethical barriers to randomization, unwillingness of funders or payers to withhold the intervention from patients (ie, the control group), or anticipated contamination of the intervention into the control group (eg, provider practice change interventions). In addition, it may be impossible to conduct an RCT because the investigator does not have control over the design of an intervention or because they are studying an event, such as a pandemic.

In the June 2020 issue of the Journal of Hospital Medicine, Coon et al1 use a type of quasi-experimental design (QED)—specifically, the interrupted time series (ITS)—to examine the impact of the adoption of ward-based high-flow nasal cannula protocols on intensive care unit (ICU) admission for bronchiolitis of heated high-flow nasal cannula protocols across hospitals. In this methodologic progress note, we discuss QEDs for evaluating the impact of healthcare interventions or events and focus on ITS, one of the strongest QEDs.

WHAT IS AN INTERRUPTED TIME SERIES?

Interrupted time series designs use repeated observations of an outcome over time. This method then divides, or “interrupts,” the series of data into two time periods: before the intervention or event and after. Using data from the preintervention period, an underlying trend in the outcome is estimated and assumed to continue forward into the postintervention period to estimate what would have occurred without the intervention. Any significant change in the outcome at the beginning of the postintervention period or change in the trend in the postintervention is then attributed to the intervention.

There are several important methodologic considerations when designing an ITS study, as detailed in other review papers.2,3,7,8 An ITS design can be retrospective or prospective. It can be of a single center or include multiple sites, as in Coon et al. It can be conducted with or without a control. The inclusion of a control, when appropriately chosen, improves the strength of the study design because it can account for seasonal trends and potential confounders that vary over time. The control can be a different group of hospitals or participants that are similar but did...
Interrupted time series can be more feasible because it requires a small number of time points for data collection. This approach is simple and has lower costs. Temporal biases can be substantial. Nonequivalent groups may not be similar in terms of important covariates.

Pre/post without control group

A group receiving the intervention is compared before and after the intervention is implemented. Analysis is based on simply estimating the difference in the outcome between the pre- and the post-groups. Unlike the interrupted time series, it does not use serial measurements over time to take into account the rate of change in the outcome in the groups.

Yes

This can be more feasible because it requires a small number of time points for data collection. This approach is simple and has lower costs. Temporal biases are a major issue.

### What is a segmented regression ITS?

Segmented regression is the statistical model used to measure (a) the immediate change in the outcome (level) at the start of the intervention and (b) the change in the trend of the outcome (slope) in the postintervention period vs that in the pre-intervention period. Therefore, the intervention effect size is expressed in terms of the level change and the slope change. To function properly, the models require several repeated (eg, monthly) measurements of the outcome before and after the intervention. Some experts suggest a minimum of 4 to 12 observations, depending on a number of factors including the stability of the outcome and seasonal variations. If changes before and after more than one intervention are being examined, there should be the minimum number of observations separating them. Unlike typical regression models, time-series models can correct for autocorrelation if it is present in the data. Autocorrelation is the type of correlation that arises when data are collected over time, with those closest in time being more strongly correlated (there are also other types of autocorrelation, such as seasonal patterns). Using available statistical software, autocorrelation can be detected and, if present, it can be controlled for in the segmented regression models.

### How Are Segmented Regression Results Presented?

Coon et al present results of their ITS analysis in a panel of figures detailing each study outcome, ICU admission, ICU length of stay, total length of stay, and rates of mechanical ventilation. Each panel shows the rate of change in the outcome per season across hospitals, before and after adoption of heated high-flow nasal cannula protocols, and the level change at the time of adoption.
To further explain how segmented regression results are presented, in the Figure we detail the structure of a segmented regression figure evaluating the impact of an intervention without a control group. In addition to the regression figure, authors typically provide 95% CIs around the rates, level change, and the difference between the postintervention and preintervention periods, along with $P$ values demonstrating whether the rates, level change, and the differences between period slopes differ significantly from zero.

**WHAT ARE THE UNDERLYING ASSUMPTIONS OF THE SEGMENTED REGRESSION ITS?**

Segmented regression models assume a linear trend in the outcome. If the outcome follows a nonlinear pattern (e.g., exponential spread of a disease during a pandemic), then using different distributions in the modeling or transformations of the data may be necessary. The validity of the comparison between the pre- and postintervention groups relies on the similarity between the populations. When there is imbalance, investigators can consider matching based on important characteristics or applying risk adjustment as necessary. Another important assumption is that the outcome of interest is unchanged in the absence of the intervention. Finally, the analysis assumes that the intervention is fully implemented at the time the postintervention period begins. Often, there is a washout period during which the old approach is stopped and the new approach (the intervention) is being implemented and can easily be taken into account.

**WHAT ARE THE STRENGTHS OF THE SEGMENTED REGRESSION ITS?**

There are several strengths of the ITS analysis and segmented regression. First, this approach accounts for a possible secular trend in the outcome measure that may have been present prior to the intervention.
to the intervention. For example, investigators might conclude that a readmissions program was effective in reducing readmissions if they found that the mean readmission percentage in the period after the intervention was significantly lower than before using a simple pre/post study design. However, what if the readmission rate was already going down prior to the intervention? Using an ITS approach, they may have found that the rate of readmissions simply continued to decrease after the intervention at the same rate that it was decreasing prior to the intervention and, therefore, conclude that the intervention was not effective. Second, because the ITS approach evaluates changes in rates of an outcome at a population level, confounding by individual-level variables will not introduce serious bias unless the confounding occurred at the same time as the intervention. Third, ITS can be used to measure the unintended consequences of interventions or events, and investigators can construct separate time-series analyses for different outcomes. Fourth, ITS can be used to evaluate the impact of the intervention on subpopulations (eg, those grouped by age, sex, race) by conducting stratified analysis. Fifth, ITS provides simple and clear graphical results that can be easily understood by various audiences.

WHAT ARE THE IMPORTANT LIMITATIONS OF AN ITS?

By accounting for preintervention trends, ITS studies permit stronger causal inference than do cross-sectional or simple pre/post QEDs, but they may be prone to confounding by cointerventions or by changes in the population composition. Causal inference based on the ITS analysis is only valid to the extent to which the intervention was the only thing that changed at the point in time between the preintervention and postintervention periods. It is important for investigators to consider this in the design and discuss any coincident interventions. If there are multiple interventions over time, it is possible to account for these changes in the study design by creating multiple points of interruption provided there are sufficient measurements of the outcome between interventions. If the composition of the population changes at the same time as the intervention, this introduces bias. Changes in the ability to measure the outcome or changes to its definition also threaten the validity of the study’s inferences. Finally, it is also important to remember that when the outcome is a population-level measurement, inferences about individual-level outcomes are inappropriate due to ecological fallacies (ie, when inferences about individuals are deduced from inferences about the group to which those individuals belong). For example, Coon et al found that infants with bronchiolitis in the ward-based high-flow nasal cannula protocol group had greater ICU admission rates. It would be inappropriate to conclude that, based on this, an individual infant in a hospital on a ward-based protocol is more likely to be admitted to the ICU.

CONCLUSION

Studies evaluating interventions and events are important for informing healthcare practice, policy, and public health. While an RCT is the preferred method for such evaluations, investigators must often consider alternative study designs when an RCT is not feasible or when more real-world outcome evaluation is desired. Quasi-experimental designs are employed in studies that do not use randomization to study the impact of interventions in real-world settings, and an interrupted time series is a strong QED for the evaluation of interventions and natural experiments.

Disclosures. The authors did not receive commercial support for the submitted work. Dr Mahant holds a grant, payable to his institution, from the Canadian Institutes of Health Research, outside the scope of the submitted work.

References