Reflection on research methods: The before and after study.

Mohammad Cheik-Hussein (cheikhussein@gmail.com)
Liverpool Hospital, Sydney https://orcid.org/0000-0002-6626-5736
Ian Harris
Liverpool Hospital, Sydney
Adriane Lewin
University of New South Wales

Technical advance

Keywords: Interrupted Time Series Analysis, Historically Controlled Study, Epidemiologic Methods, Research Design

Posted Date: October 8th, 2019

DOI: https://doi.org/10.21203/rs.2.15704/v1

Abstract

Background Before and after studies allow for the investigation of population-level health interventions and are a valuable study design in situations where randomisation is not feasible. The before and after study design involves measuring an outcome both before and after an intervention and comparing the outcome rates in both time periods to determine the effectiveness of the intervention. These studies do not involve a contemporaneous control group and must therefore take into account any underlying secular trends in order to separate the effect of the intervention from any pre-existing trend. Neglecting this important step can lead to spurious results.

Methods To illustrate the importance of accounting for underlying trends, we performed a before and after study assessing 30-day mortality in hip fracture patients without any actual intervention, and instead designated an arbitrarily-chosen time-point as our ‘intervention’. We did this to ensure that we were basing our results exclusively on the underlying trend throughout the studied period and also to enable us to show that even an intervention of nothing may be spuriously interpreted to have an effect if the before and after study is incorrectly analysed.

Results We found a secular trend in our data showing improving 30-day mortality in hip fracture patients in our institution. We then demonstrated that disregarding this underlying trend showed that our intervention of nothing ‘resulted’ in a significant decrease in mortality, from 6.7% in the ‘before’ period to 3.1% in the ‘after’ period (p<0.0008). This apparent impact on mortality disappeared when we accounted for the underlying trend in our analysis (IRR of 0.75, 95% CI 0.32 – 1.78; p=0.5). In the context of declining 30-day mortality following hip fracture, failure to consider the existing underlying trend lead us to believe that it was our ‘intervention’ that ‘caused’ the decrease in mortality in the ’after’ period compared to the ‘before’ period when our results clearly show that mortality was decreasing irrespective of any intervention.

Conclusion Our study highlights the importance of appropriate measurement and consideration of underlying trends when analysing data from before and after studies and illustrates what can happen should researchers neglect this important step.

Key Messages

Before and after studies are valuable in situations where randomisation is not feasible or ethical.

To be able to assess the effectiveness of any intervention, before and after studies must measure the underlying trend existing in the study population prior to the implementation of the intervention.

To avoid spurious results, data analyses must adjust for any underlying trend in order to measure the true intervention effect.

Existing publications often neglect these robust and appropriate analyses, possibly leading to spurious or inaccurate conclusions.

Interrupted Time Series Analysis, Historically Controlled Study, Epidemiologic Methods, Research Design

Introduction

What is a before and after study?

Before and after studies compare group-level data for a specified population prior to and following an intervention and assess changes in outcomes between the two periods. These studies may also be referred to as interrupted time series, pre-post studies or historically controlled studies and are often used in situations where resources are limited or where randomization is either not feasible or not ethical. (1) Compared with other study designs, before and after studies can be relatively simple to implement, less logistically demanding (particularly if they are observing changes in standard practice) and cost-effective. (2)

Examples of before and after studies

Before and after studies allow researchers to answer questions such as: Does the opening of a new trauma centre reduce admissions in surrounding hospitals? (3) Do sugar taxes reduce the consumption of sweetened beverages? (4) Does the introduction of compulsory helmet legislation for cyclists reduce the rate of head injury? (5) These and other questions that may be difficult or impossible to investigate with a randomised controlled trial lend themselves easily to the before and after study design.

Pitfalls and dangers of before and after studies

As with all study designs, however, before and after studies have important limitations. Many health conditions or states improve over time, due either to the natural history of the condition or to concomitant interventions that are not considered or measured as part of a study, and it can be difficult to isolate the effect of a single intervention from the myriad other changes occurring around us all the time, in the health system or in society at large.(6) For example, European survival rates for breast cancer have been improving since the 1980s due to a combination of improvements in therapeutics, early diagnosis and by the improvement in engagement in screening programs (7). A before and after study measuring the effectiveness of an intervention on survival rates of European breast cancer patients must therefore identify and account for this trend to avoid spuriously finding a benefit of the study intervention.

A failure to measure and account for these underlying secular trends may cause the studied intervention to appear, spuriously, to be effective. This can contribute to the relative overestimation of effectiveness that has been observed when using a before and after study design compared to randomised trials. (8) A careful analysis of data from a before and after study, therefore, should focus on assessing whether an intervention has an effect *beyond that due to any pre-existing trends.* (9)

Why is this important?

Prior to performing our own before and after study, we reviewed a sample of 249 randomly selected before and after studies of population-level interventions and found that 77% of these studies reported that the intervention under assessment was effective. Without accounting for underlying trends in the data (by performing unadjusted, simple before-and-after comparisons) is it likely that many of the improvements seen were spurious findings. The objective of this article is therefore to highlight the importance of identifying and adjusting for underlying trends in before and after studies, and to show what can happen when this step is omitted. To illustrate this, we used data from our institution to report 30-day hip fracture mortality before and after *nothing,* i.e., we assigned an arbitrary mid-point date as the 'intervention' with no specific intervention having taken place. We then compared the 30-day mortality in the before and after periods in two ways: first, without considering the underlying trend; then, using an approach described previously by Bernal et al., (10) by taking into account the underlying secular trend.

Methods And Results

Using hip fracture mortality as an example.

This before and after study investigated changes in 30-day mortality at a single tertiary care institution from 2010 to 2016 (inclusive) using two different analytical approaches: one which accounted for any trend during the 'before' period and one which did not. For illustrative purposes, we arbitrarily chose 1st January 2014 to denote the date of implementation of an 'intervention'; in reality, there was no intervention. We did this in order to show that even an intervention of nothing may be spuriously found to have an effect if an underlying trend in the data already exists. The study used de-identified routinely collected departmental audit data.

This study included patients aged 50 years or older who presented to Liverpool Hospital, New South Wales, Australia between January 2010 and December 2016 for treatment of hip fractures. The outcome of interest, 30-day mortality, is routinely audited by our institution as part of standard care using direct patient or family contact.

We used a pre-specified significance level of $p<0.05$ and performed all analyses using Stata 15.1 (www.stata.com; StataCorp; College Station, TX).

Analysis of data: two ways

Analysis 1: Not accounting for underlying trend

For the first analysis, we used a chi-square test to compare the proportion of people who died within 30 days of a hip fracture between two time periods, before (2010–2013), and after (2014–2016) the 'intervention'.

This analytical approach, which did not incorporate or even consider any pre-existing trend in 30-day mortality, found a statistically significant decrease in 30-day mortality between the ‘before’ and ‘after’ periods; 65 (6.7%) of 970 hip fracture patients died within 30 days of presentation during the ‘before’ period (2010–2013), compared to 24 (3.1%) of 765 in the ‘after’ period (2014–2016; p = 0.0008) (Figure 1). Interpretation: January 1st, 2014 was such an important day that it set into motion a decrease in hip fracture mortality.

[Figure 1]

Analysis 2: Accounting for underlying trend

We found markedly different results when we performed an interrupted time series analysis as described by Bernal et al. (10) Using a generalised linear model specifying a Poisson distribution, we measured the underlying 30-day mortality during the ‘before’ period, then projected that measure into the ‘after’ period. This projection, called a counterfactual, shows what we would expect to see if the pre-existing trend in 30-day hip fracture mortality during the ‘before’ period were to continue unchanged into the ‘after’ period i.e., beyond January 1, 2014. Modelling allowed us to compare the two time periods (‘before’ i.e., 2010–2013 vs ‘after’ i.e., 2014–2016) for any difference in outcome while adjusting for any underlying trend in 30-day mortality.

Using this analytical approach, we found that our intervention had no effect, i.e., there was a non-significant decrease in 30-day mortality during the ‘after’ period (2014–2016) compared to the ‘before’ period (2010–2013), with an incidence rate ratio [IRR] of 0.75 (95% confidence interval 0.32—1.78; p = 0.5) (Figure 2).

[Figure 2]

Discussion

When applying a statistical method that was simple but inappropriate for this context i.e., not considering the possibility of any underlying secular trend, we found a statistically significant decline in 30-day mortality when comparing periods before and after January 1st, 2014, an arbitrarily chosen time point denoting no specific or actual intervention. Following adjustment for the underlying trend in 30-day mortality using an appropriate analytical method, however, the ‘intervention’ was not effective, and the difference between periods was not significant.

This illustrates the importance of performing statistical analyses that consider pre-existing trends when assessing the possible effectiveness of an intervention in before and after studies to prevent spurious results. Using the first analytical approach, any intervention could have been deemed to be effective, regardless of its true effectiveness, as in the environment of decreasing 30-day hip fracture mortality, the effect of this pre-existing trend could be mistakenly attributed to the ‘intervention’.

When considering possible drivers of 30-day mortality in our patients, we reflected on clinical practice during this period. Overall practice did not change in our institution over the period under study. No policy changes were introduced regarding the timing or urgency of hip fracture surgery and there was no change in the involvement of the orthogeriatric service, a factor that has been previously shown to influence 30-day mortality in our state. (11) We discussed 30-day mortality at our institution with staff, who made several suggestions to explain the reduction in mortality over the studied time period. These included an increased awareness in the problems facing hip fracture patients, the addition of two new consultant orthopaedic staff, the addition of a second Clinical Nurse Consultant to the ward, the introduction of regular Structured Interdisciplinary Bedside Rounding, and the initiation of data collection for the Australia & New Zealand Hip Fracture Registry. In the context of a gradual decline in mortality over time, it is likely that a simply-analysed before-and-after study of nearly any intervention, whether or not that intervention was actually effective, would have reported a significant decline in mortality.

The wide range of mortality for each year (from 1.4% to 8.1%) and the decline over time is consistent with findings in similar published studies. A 2016 review of 16 separate international studies including more than a million elderly patients who experienced traumatic hip fractures between 1986 and 2013 reported 30-day mortality varying from 1.4% to 10% with multiple factors contributing to a steady reduction in 30-day mortality from year to year including improved time to surgery, multidisciplinary care and implementation of NICE guidelines. (12)

We have shown that, even in the absence of any specific intervention, it is possible to find a spurious statistically significant result using a simple comparison of outcomes between two discrete time periods in a before and after study in the presence of an underlying trend. Using an appropriate analytical method to measure and account for the underlying trend permits appropriate comparison of two periods in studies of this type. Our findings highlight the importance of adjusting for underlying secular trends and recommend a cautious interpretation when evaluating the effects of interventions of before and after studies.

Declarations

- Ethics approval and consent to participate: Our study is a secondary analysis on routinely collected data. Ethics approval was not sought.
- Consent for publication: Not applicable
- Availability of data and materials: Data used in this study is routinely collected in our institution and collated through a national registry.
- Competing interests: The authors declare that they have no competing interests.
- Funding: Not applicable
- Authors’ contributions: MC contributed significantly to preparation of the manuscript and interpretation of results. IH contributed significantly to inception of the study and preparation of the

manuscript. AL contributed significantly to data analysis and preparation of the manuscript. All authors read and approved the final manuscript.

- Acknowledgements: Not applicable

References

1.Gordis L. Epidemiology. Fifth edition. ed. Philadelphia, PA: Elsevier/Saunders; 2014. xv, 392 pages p.

2.Kontopantelis E, Doran T, Springate DA, Buchan I, Reeves D. Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis. BMJ. 2015;350:h2750.

3.Martin CL, Aldridge PJ, Harris AM, Perkins CH. Opening a New Level II Trauma Center Near an Established Level I Trauma Center: Is This Good for Trauma Care? J Orthop Trauma. 2016;30(10):517–23.

4.Silver LD, Ng SW, Ryan-Ibarra S, Taillie LS, Induni M, Miles DR, et al. Changes in prices, sales, consumer spending, and beverage consumption one year after a tax on sugar-sweetened beverages in Berkeley, California, US: A before-and-after study. PLoS Med. 2017;14(4):e1002283.

5.Walter SR, Olivier J, Churches T, Grzebieta R. The impact of compulsory cycle helmet legislation on cyclist head injuries in New South Wales, Australia. Accid Anal Prev. 2011;43(6):2064–71.

6.Penfold RB, Zhang F. Use of interrupted time series analysis in evaluating health care quality improvements. Acad Pediatr. 2013;13(6 Suppl):S38–44.

7.Sant M, Francisci S, Capocaccia R, Verdecchia A, Allemani C, Berrino F. Time trends of breast cancer survival in Europe in relation to incidence and mortality. Int J Cancer. 2006;119(10):2417–22.

8.Sacks H, Chalmers TC, Smith H, Jr. Randomized versus historical controls for clinical trials. Am J Med. 1982;72(2):233–40.

9.Polus S, Pieper D, Burns J, Fretheim A, Ramsay C, Higgins JPT, et al. Heterogeneity in application, design, and analysis characteristics was found for controlled before-after and interrupted time series studies included in Cochrane reviews. J Clin Epidemiol. 2017;91:56–69.

10.Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol. 2017;46(1):348–55.

11.Zeltzer J, Mitchell RJ, Toson B, Harris IA, Ahmad L, Close J. Orthogeriatric services associated with lower 30-day mortality for older patients who undergo surgery for hip fracture. Med J Aust. 2014;201(7):409–11.

12.Giannoulis D, Calori GM, Giannoudis PV. Thirty-day mortality after hip fractures: has anything changed? Eur J Orthop Surg Traumatol. 2016;26(4):365–70.

Figures

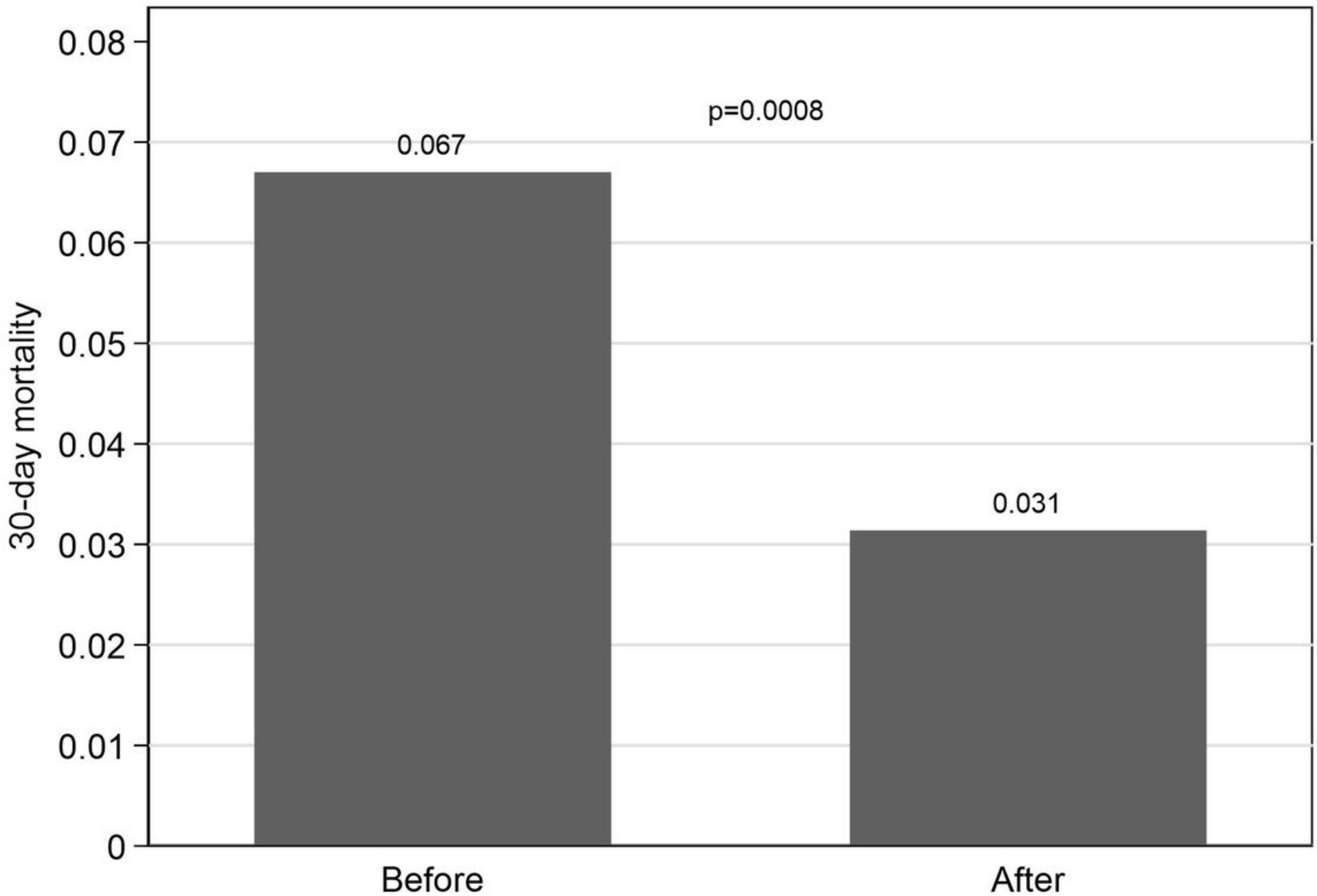

Figure 1

Proportion of people who died within 30 days of hip fracture during the ‘before’ (2010-2013) and ‘after’ (2014-2016) time periods.

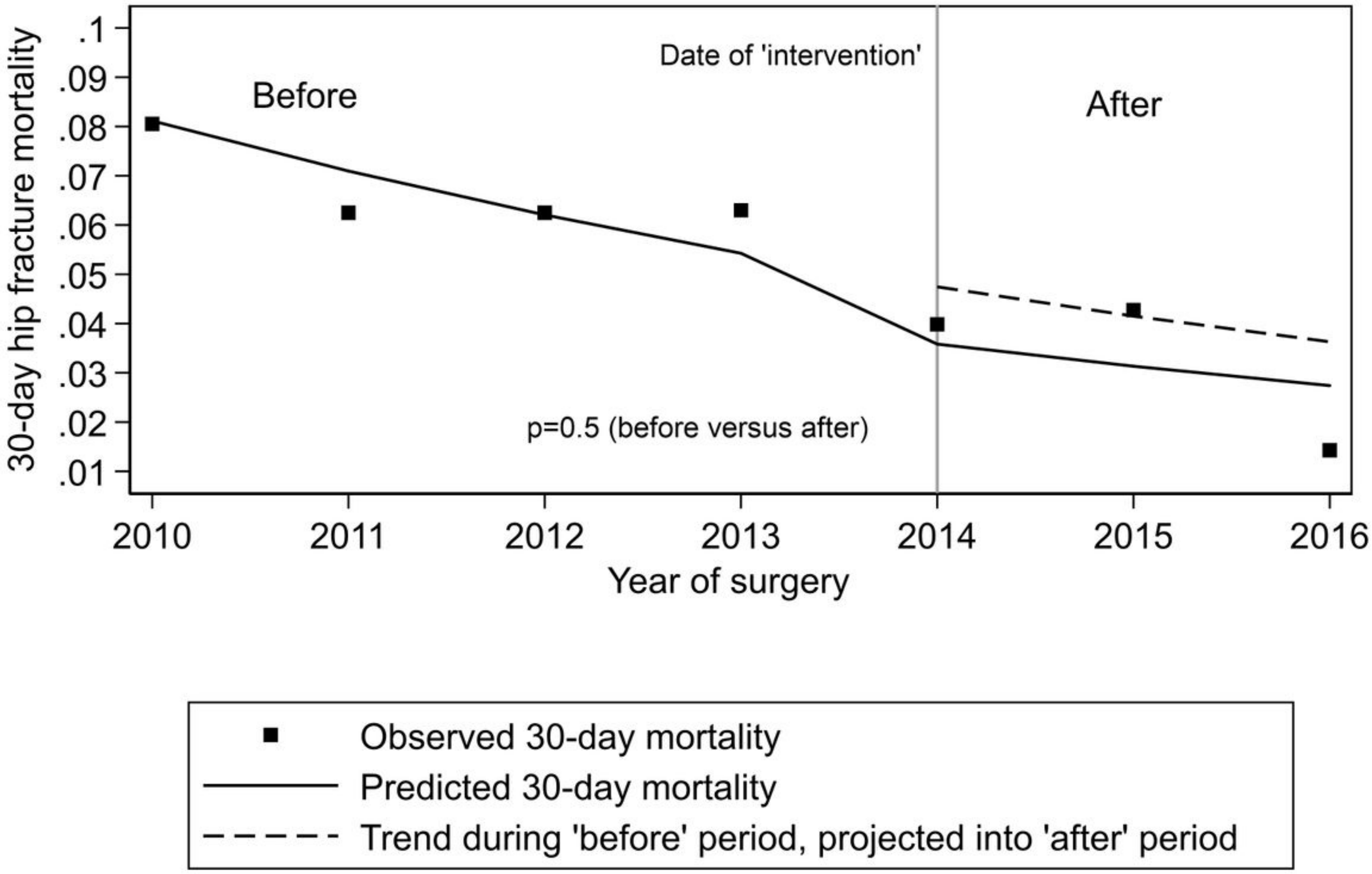

Figure 2

30-day hip fracture mortality before and after an ‘intervention’. The dotted line projects the 30-day mortality during the ‘before’ period into the ‘after’ period.