Epidemiology and Infection

cambridge.org/hyg

Letter to the Editor

Cite this article: Zhao S (2020). The insignificant structural break in the relationship between improved observed hand hygiene on methicillin-resistant *Staphylococcus aureus* bloodstream infection rates in Ireland: a data-driven re-analysis. *Epidemiology and Infection* **148**, e297, 1–2. https://doi.org/10.1017/S0950268820002733

Received: 6 July 2020
Revised: 28 September 2020
Accepted: 30 September 2020

Author for correspondence:
Shi Zhao, E-mail: zhaoshi.cmsa@gmail.com

CAMBRIDGE UNIVERSITY PRESS

The insignificant structural break in the relationship between improved observed hand hygiene on methicillin-resistant *Staphylococcus aureus* bloodstream infection rates in Ireland: a data-driven re-analysis

Shi Zhao[1,2]

[1]JC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong, China and [2]CUHK Shenzhen Research Institute, Shenzhen, China

To the Editor

Segmented regression models are commonly used for 'before-and-after' comparison in interventional program evaluation [1]. Smiddy *et al.* [2] analysed the bloodstream infection (BSI) rate time series in exploring the change in their temporal trends before and after the implementation of observational hand hygiene auditing in Ireland. They reported that the decreasing trends of methicillin-resistant *Staphylococcus aureus* (MRSA) BSI rate significantly slowed, with *P*-value = 0.007, from 5% to 2% decrease per quarter pre- and post-intervention, respectively, see their Table 2. In other words, they found that the structural break, i.e. the change in slope in [2], in the decreasing trends of MRSA BSI rate is statistically significant. However, by observing the temporal trends presented in their Figure 2 (the panel at row 1 and column 2), I considered this structural break is not obvious, and it may not reach statistical significance. In this study, I re-analysed the data in [2] to examine the likelihood of the structural break in the decreasing trends of MRSA BSI rate.

The raw data in [2] were unavailable, and thus I digitised the panel at row 1 and column 2 in Figure 2 of their work using WebPlotDigitizer (version 4.2, https://automeris.io/WebPlotDigitizer). I considered two regression models as follows. They were

- the baseline model: $\boldsymbol{E}[\ln(\text{BSI}_t)] = \alpha_0 + \alpha_1 t + \alpha_2\text{Gap}$, and
- the full model: $\boldsymbol{E}[\ln(\text{BSI}_t)] = \beta_0 + \beta_1 t_1 + \beta_2 t_2 + \beta_3\text{Gap}$, which was the segmented regression as the same as in [2].

The $\boldsymbol{E}[\cdot]$ denoted the expectation and the BSI_t denoted the MRSA BSI rate at the *t*th time interval. The settings, meaning and interpretation of other notations were exactly the same as in [2], but the fitting was conducted with full likelihood frameworks. In this re-analysis, I examined the following three aspects relevant to the estimate of a structural break in BSI_t. They included

- the consistency, i.e. the effect size and statistical significance, between the β_1 and β_2 estimates in this study and those in [2],
- the consistency between the statistical significance of the change in slope, i.e. $(\beta_1 - \beta_2)$, in this study and that in [2], and
- the statistical significance of the likelihood-ratio (LR) test of the full model against the baseline model.

All analyses were carried out in **R** software (version 3.6.0) [3].

I reported that the β_1 and β_2 estimates were consistent with those in Table 2 of [2]. Different from the significant level estimated in [2], I found the *P*-value = 0.088 appears not statistically significant at the 5% level, against *P*-value = 0.007 < 0.05 in [2], for the change in slope, which indicated the difference in slopes was statistically unclear. Moreover, I noticed that the 95% confidence intervals (95%CI) of $\exp(\beta_1)$ (from 0.93 to 0.97) and $\exp(\beta_2)$ (from 0.97 to 0.99) estimates in [2] were extremely close. This might also be a sign of no difference in β_1 and β_2, and thus a structural break may not have occurred. For the LR test, the *P*-value = 0.066, which implied the structural break unlikely occurred. As shown in Figure 1 (of this study), the fitting results of the baseline and full models almost overlapped.

I note that different fitting framework, as well as, the testing methods were adopted in this study and in [2]. The likelihood inference with LR test was used in this study, whereas the robust estimation with Wald test was carried out to obtain an estimate of the variance (Huber-White) for the data as a whole in [2]. However, as pointed out in [4],

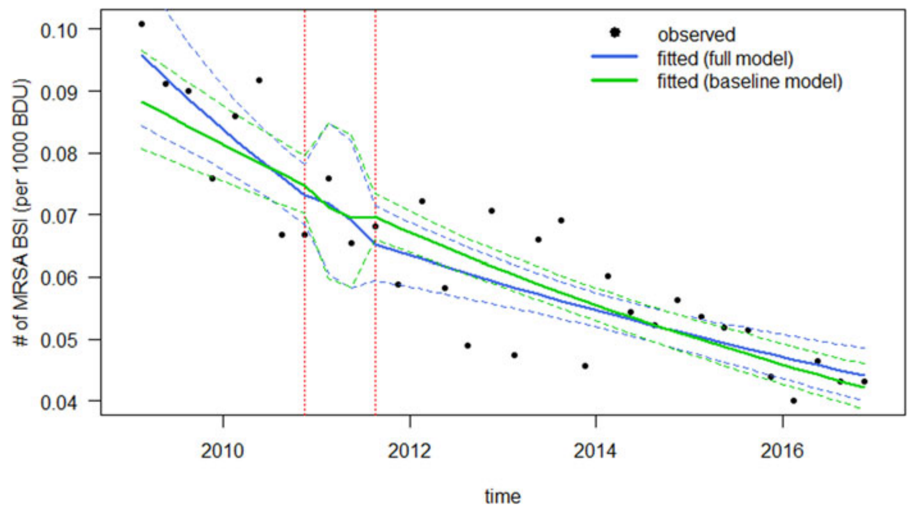

Fig. 1. The fitting results of the association between MRSA BSI incidence per 1000 bed days used (BDU) in Ireland from 2009 to 2016. The black dots are the observed data in Figure 2 (the panel at row 1 and column 2) of [2]. The blue curves are the fitting results from the full model and the green curves are the fitting results from the baseline model. Bold curves are the mean fitting results and the dashed curves are the 95%CI.

> 'white robust standard errors are universally used in econometrics, (and) their finite sample properties lead to over-rejection under the null hypothesis, sometimes by a large amount.'

This difference in thecourse of analysis probably leads to the inconsistency in the estimates or testing outcomes.

Although this inconsistency between Smiddy *et al.*'s findings and the findings in this study was unlikely to affect the main conclusions in [2], the inconsistency in analysis outcomes from different model setting or frameworks should be considered with caution.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0950268820002733.

Acknowledgements. I acknowledge that reference [4] was initially mentioned by Professor Gabrielle Kelly, who provided statistical adjudication, during the reviewing process of this work.

Author contributions. SZ conceived the study, carried out the analysis, discussed the results, drafted the first manuscript, critically read and revised the manuscript, and gave final approval for publication.

Financial support. This work was not funded.

Consent for publication. Not applicable.

Disclaimer. The funding agencies had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript or decision to submit the manuscript for publication.

Conflict of interests. The author declared no competing interests.

Ethical standards. Ethical approval or individual consent was not applicable.

Availability of data and materials. The raw data digitised from the panel at row 1 and column 2 in Figure 2 of [2], and **R** codes were both available via the online supplementary files.

References

1. **Linden A and Adams JL** (2011) Applying a propensity score-based weighting model to interrupted time series data: improving causal inference in programme evaluation. *Journal of Evaluation in Clinical Practice* **17**, 1231–1238.
2. **Smiddy MP *et al.*** (2020) Impact of improved observed hand hygiene on bloodstream infection rates in Ireland. A prospective segmented regression analysis, 2009–2016. *Epidemiology and Infection* **148**, e83.
3. **Team RC** (2013) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: Team RC.
4. **Hausman J and Palmer C** (2012) Heteroskedasticity-robust inference in finite samples. *Economics Letters* **116**, 232–235.