



# Counterpoint: The value of benchmarking in observational studies of the longer term safety of antenatal corticosteroids

Jennifer A. Hutcheon | Jessica Liauw

Department of Obstetrics and Gynaecology, University of British Columbia, British Columbia, Vancouver, Canada

## Correspondence

Jennifer A. Hutcheon, Department of Obstetrics and Gynaecology, BC Children's & Women's Hospital, Vancouver, Canada.

Email: [jhutcheon@bcchr.ca](mailto:jhutcheon@bcchr.ca)

## Funding information

Canadian Institutes of Health Research

We appreciate the opportunity to respond to Caniglia<sup>1</sup> and Kajantie's<sup>2</sup> commentaries on our recent work<sup>3</sup> examining the generalisability of the Antenatal Late Preterm Steroids (ALPS) trial<sup>4</sup> of antenatal corticosteroid administration at late preterm gestation (34–36 weeks) to real-world populations.

We agree with Kajantie<sup>2</sup> that the unknown potential for harm to child neurodevelopment is a critical knowledge gap for informing clinical decision-making on antenatal corticosteroid administration at late preterm ages. We also agree that observational data play an important role in understanding longer-term safety, as evidence from follow-up of randomised trials is unconvincing due to small sample sizes and high losses to follow-up. Yet, we have profound concerns that many existing observational studies assessing the safety of antenatal corticosteroid for child neurodevelopment, including the study from Finland highlighted by Kajantie,<sup>5</sup> demonstrate findings that reflect confounding, rather than a true effect of the medication. Here, we outline the theoretical potential for this bias and use a whole-population cohort from British Columbia, Canada to demonstrate how the strategy proposed by Caniglia,<sup>1</sup> benchmarking, provides quantitative evidence supporting this concern.

As with most population database studies, many potential confounders in the Finnish study<sup>5</sup> were documented only as the presence or absence of a diagnosis (e.g. hypertensive disorder of pregnancy yes/no). Clinically, however, these conditions present with a range of severity. For example, the international classification of diseases (ICD) codes for hypertensive disorders of pregnancy O10 and O13–O15 run the gamut from mild preeclampsia to life-threatening eclampsia. As more severe presentations of a condition will be more likely to lead to a preterm delivery than milder presentations, more severe presentations will also be more likely to be administered antenatal steroids. Thus, even after covariate control

for the presence of a given diagnosis (i.e. within strata of a diagnosis), pregnancies exposed to antenatal steroids will likely still be a systematically higher risk group than unexposed pregnancies. As more severe disease presentation is likely also linked with greater risk of adverse child neurodevelopment (as previously shown for hypertension<sup>6</sup>), the differences in clinical risk profile between exposure groups likely creates a spurious association between antenatal corticosteroid administration and longer term adverse child health (i.e. confounding).

The sibship analysis in the Finnish study<sup>5</sup> does little to alleviate these concerns. Sibling analyses can be a powerful way to control for difficult-to-measure confounders that remain relatively constant between an individual's pregnancies, such as genetics, socio-economic status and lifestyle-related risk factors. In the context of antenatal corticosteroids and child neurodevelopment, however, these factors are unlikely to be meaningful confounders. To introduce confounding, the factor must be a determinant of antenatal corticosteroid administration. Clinical decision-making on antenatal corticosteroid administration, however, is unlikely to be influenced by factors such as genetics or lifestyle. Rather, the reason why an individual receives antenatal steroids in one pregnancy and not another (i.e. how the discordant exposure status between pregnancies needed for sibship analyses arises) is more likely to be influenced by pregnancy-specific complications that cause the clinician to anticipate an imminent preterm birth (such as severe preeclampsia). The sibship design, however, does not control for factors that differ between an individual's pregnancies. As a result, the findings of the sibship analysis are likely also confounded by clinical risk profile, producing the same spurious associations described above.

The use of benchmarking proposed by Caniglia and Hinkle<sup>1</sup> provides an excellent strategy to assess concerns of confounding

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Paediatric and Perinatal Epidemiology* published by John Wiley & Sons Ltd.

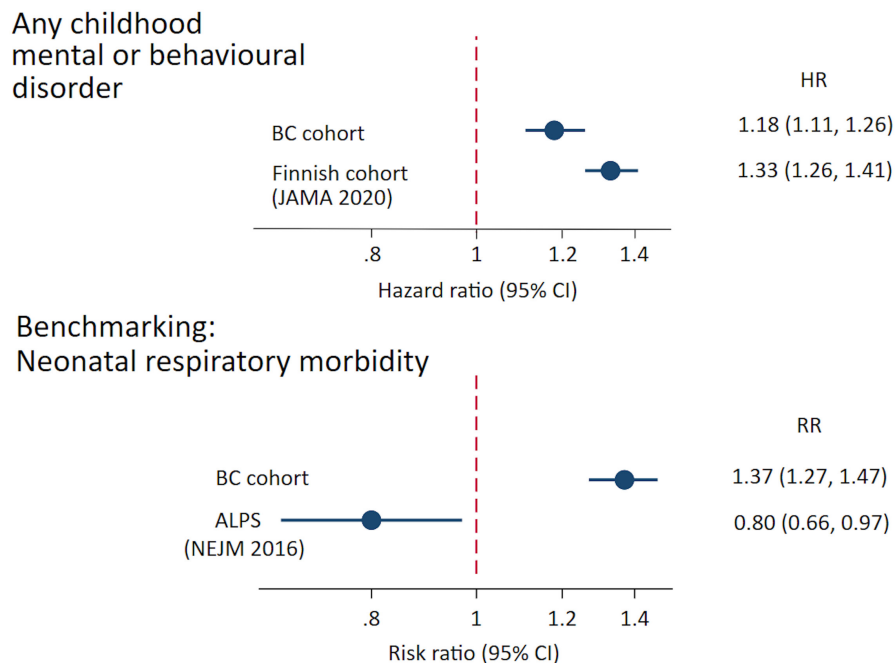
in observational studies of the longer term safety of antenatal corticosteroids. If researchers can replicate the protective effect of antenatal steroids on neonatal respiratory morbidity reported in randomised trials using their observational design, this would greatly increase confidence that their study has adequately controlled for confounding when studying other longer term outcomes. Alternatively, if the observational design finds that antenatal corticosteroid administration is associated with an increased risk of neonatal respiratory morbidity or has no association with neonatal respiratory morbidity, this suggests that findings for longer term safety outcomes may also be confounded.

We sought to demonstrate how the broad principles of benchmarking could help to assess the potential for confounding in observational studies of the association between antenatal corticosteroid administration and child neurodevelopment. We used a cohort highly similar to that used in the aforementioned Finnish study<sup>5</sup>: a large birth registry cohort from British Columbia, Canada, 2000–2013 ( $n = 526,525$ ),<sup>7</sup> linked with population-based follow-up using child hospitalisation<sup>8</sup> and outpatient physician visit records<sup>9</sup> to 2018. Further details on the cohort (including ethical approval) have previously been published.<sup>10,11</sup> The list of confounders available in this cohort was highly similar to that in the Finnish study (exceptions noted in Figure 1 footnote). We limited our cohort to 368,825 births with available pre-pregnancy body mass index (BMI), but note that our results were unchanged when repeating analysis in the full cohort without controlling for BMI. We used the same primary

outcome definition as the Finnish study of 'any childhood mental and behavioural disorder' (ICD-10 codes F00–F99), and a similar analytic approach.

The estimated association between antenatal corticosteroid administration and any childhood mental or behavioural disorder in our cohort was reasonably comparable to that reported in the Finnish study<sup>5</sup>: a hazard ratio of 1.18 and 1.33, respectively (Figure 1, top panel). We next examined the extent to which benchmarking could assess the potential for confounding in these analyses. We built a logistic regression model using the same population and the same covariates, but with an outcome of neonatal respiratory morbidity/mortality (ICD-10 code P22 or in-hospital newborn death). This model estimated that antenatal corticosteroid administration was associated with 37% increased odds of neonatal respiratory morbidity/mortality (Figure 1, lower panel). Given that randomised trial evidence has shown that antenatal corticosteroids decrease, not increase, the risk of neonatal respiratory morbidity and mortality,<sup>4</sup> our findings suggest that the covariate adjustment for a broad list of risk factors was not sufficient to remove confounding by differences in clinical risk profile between those who do and do not receive steroids.

These analyses highlight a critical challenge in benchmarking: the benefits of benchmarking outlined by Caniglia<sup>1</sup> are contingent on being able to replicate randomised trial findings using observational designs. Their phrase 'If researchers can successfully replicate the ALPS trial findings...' should not be taken lightly. There is a very good



**FIGURE 1** Use of benchmarking to estimate associations between antenatal corticosteroid administration and child health. Risk ratios and hazard ratios in British Columbia cohort ( $n = 368,825$ , 2000–2013) were adjusted for child birth year, sex, 5-minute Apgar score, admission to a neonatal intensive care unit, weight and gestational age at birth, and maternal age at delivery, parity, mode of delivery, smoking during pregnancy, prepregnancy body mass index (BMI), premature rupture of membranes (ICD-10 code O42), gestational diabetes (ICD-10 code O24), and hypertension in pregnancy (ICD-10 codes O10, O13–O15), maternal lifetime mental disorder diagnosis. Information on 1-minute Apgar score was not available, and maternal mental health disorder was based only on information collected during pregnancy and delivery (not lifetime). ALPS, Antenatal Late Preterm Steroids; BC, British Columbia; HR, hazard ratio; RR, risk ratio

reason why randomisation is viewed as the most reliable method to prevent confounding in studies of medical interventions, and why confounding by indication is a serious concern in observational studies.<sup>12</sup> Even with careful emulation of the target trial, standard observational analyses may often be unable to replicate randomised trial findings. Rather, we propose that quasi-experimental designs, such as difference-in-differences, instrumental variable and regression discontinuity, should be given greater consideration when seeking to replicate perinatal randomised trial findings using observational data.<sup>13</sup> By identifying situations in which treatment allocation is made through mechanisms unrelated to individual-level clinical risk profile, these quasi-experimental designs are likely much better able to control for confounding by indication than standard covariate adjustment.

We recently used a regression discontinuity design<sup>13</sup> to examine the longer term safety of antenatal corticosteroids for child neurodevelopment.<sup>10</sup> This design took advantage of the fact that, in Canada, clinical practice guidelines recommended antenatal corticosteroid administration for individuals presenting with imminent preterm birth up to 33 weeks and 6 days (33+ 6 weeks), but not one day later, at 34+ 0 weeks. As foetuses immediately on either side of this cut-off are reasonably similar in their clinical risk profiles but have very different levels of exposure to antenatal corticosteroids, this 'natural randomization' could be exploited to gain an understanding of the consequences of corticosteroid exposure on child health independent of clinical risk profile. Using the principals of benchmarking, we first demonstrated that the regression discontinuity design produces estimates of the effect of antenatal corticosteroids on neonatal respiratory morbidity/mortality that is highly comparable to those from randomised trials: a risk ratio for antenatal corticosteroids of 0.81 (95% CI 0.72, 0.91) in our design<sup>11</sup> compared with 0.80 (95% CI 0.66, 0.97) for the ALPS primary outcome of respiratory morbidity.<sup>4</sup>

After confirming compatibility with randomised trial findings for neonatal outcomes, we extended the design to examine longer term child neurodevelopment outcomes, and found no adverse consequences of routine antenatal corticosteroids administration for kindergarten-aged child development test scores<sup>11</sup> or risk of attention-deficit hyperactivity disorder.<sup>12</sup> In contrast, the Finnish study reported an increased risk of adverse neurodevelopmental outcomes,<sup>5</sup> however these conclusions relied on a study design that produced results incompatible with randomised trial findings for neonatal outcomes. Taken together, these analyses demonstrate how benchmarking can help to make sense of conflicting research findings in this field and ultimately to inform clinical care: authors of clinical practice guidelines should look for benchmarking to help determine which observational evidence is most suitable to guide patient counselling.

In conclusion, we echo a common theme raised in both commentaries: the need to make better use of observational data to optimise the translation of randomised trial evidence into clinical practice. We propose that benchmarking should be a necessary step for studies seeking to investigate the longer term safety of perinatal interventions

using observational designs, and encourage perinatal researchers to make better use of quasi-experimental designs towards this end.

## ACKNOWLEDGEMENTS

The authors thank Tim Choi from Population Data BC, Vancouver, Canada for his invaluable assistance in data access. All inferences, opinions and conclusions drawn in this study are those of the authors, and do not reflect the opinions or policies of the Data Stewards.

## AUTHOR CONTRIBUTION

JAH and JL co-conceived the idea for the Commentary. JAH wrote the first draft of the manuscript, conducted statistical analyses, and produced the Figure. JL made scientific and editorial revisions to manuscript.

## FUNDING INFORMATION

This work was funded by the Canadian Institutes of Health Research. JAH holds a Canada Research Chair in Perinatal Population Health. JL holds a Michael Smith Foundation for Health Research Health Professional Investigator award.

## CONFLICT OF INTEREST

None to declare.

## DATA AVAILABILITY STATEMENT

The data used in this study cannot be shared by the research team under the conditions through which they were made available. However, the data can be accessed by submitting a data request to PopData BC ([https://www.popdata.bc.ca/data\\_access/DAR\\_process](https://www.popdata.bc.ca/data_access/DAR_process)).

## ORCID

Jennifer A. Hutcheon  <https://orcid.org/0000-0003-0502-3423>

## REFERENCES

1. Caniglia EC, Hinkle SN. Point: Benchmarking can supplement transportability to answer critical questions about the effectiveness of antenatal corticosteroid administration. *Paediatr Perinat Epidemiol*. 2022. doi:10.1111/ppe.12923
2. Kajantie E. Antenatal corticosteroid preterm in the late preterm period- are the benefits worth the potential risks? *Paediatr Perinat Epidemiol*. 2022.
3. Hutcheon JA, Liauw J. Improving the external validity of Antenatal Late Preterm Steroids trial findings. *Paediatr Perinat Epidemiol*. 2022. doi:10.1111/ppe.12856
4. Gyamfi-Bannerman C, Thom EA, Blackwell SC, et al. Antenatal betamethasone for women at risk for late preterm delivery. *N Engl J Med*. 2016;374(14):1311-1320.
5. Raikonen K, Gissler M, Kajantie E. Associations between maternal antenatal corticosteroid treatment and mental and behavioral disorders in children. *JAMA*. 2020;323(19):1924-1933.
6. Wang H, Laszlo KD, Gissler M, et al. Maternal hypertensive disorders and neurodevelopmental disorders in offspring: a population-based cohort in two Nordic countries. *Eur J Epidemiol*. 2021;36(5):519-530.

7. Perinatal Services BC [creator]. 2018 British Columbia Data Registry. Population Data BC. Data Extract. PSBC. <http://www.perinataleservicesbc.ca/health-professionals/data-surveillance/perinatal-data-registry>
8. British Columbia Ministry of Health [creator]. 2018 Medical Services Plan (MSP) Payment Information File. Population Data BC. Data extract. BC Ministry of Health. [www.popdata.bc.ca/data](http://www.popdata.bc.ca/data)
9. British Columbia Ministry of Health [creator]. 2018 Discharge Abstract Database (Hospital Separations). Population Data BC. Data extract. BC Ministry of Health. [www.popdata.bc.ca/data](http://www.popdata.bc.ca/data)
10. Hutcheon JA, Harper S, Liauw J, Skoll MA, Srour M, Strumpf EC. Antenatal corticosteroid administration and early school age child development: a regression discontinuity study in British Columbia, Canada. *PLoS Med.* 2020;17(12):e1003435.
11. Hutcheon JA, Strumpf EC, Liauw J, et al. Antenatal corticosteroid administration and attention-deficit/hyperactivity disorder in childhood: a regression discontinuity study. *CMAJ.* 2022;194(7):E235-E241.
12. Zhang X, Faries DE, Li H, Stamey JD, Imbens GW. Addressing unmeasured confounding in comparative observational research. *Pharmacoepidemiol Drug Saf.* 2018;27(4):373-382.
13. Barnighausen T, Oldenburg C, Tugwell P, et al. Quasi-experimental study designs series-paper 7: assessing the assumptions. *J Clin Epidemiol.* 2017;89:53-66.

**How to cite this article:** Hutcheon JA, Liauw J. Counterpoint: The value of benchmarking in observational studies of the longer term safety of antenatal corticosteroids. *Paediatr Perinat Epidemiol.* 2023;37:15-18. doi: [10.1111/ppe.12937](https://doi.org/10.1111/ppe.12937)