PRENATAL PARACETAMOL EXPOSURE AND WHEEZING IN CHILDHOOD: AN APPLICATION OF THE TARGETED MAXIMUM LIKELIHOOD ESTIMATION

Moccia Chiara[1], Popovic Maja[1], Moirano Giovenale[1], Pizzi Costanza[1], Zugna Daniela[1], Richiardi Lorenzo[1], Maule Milena[1]

Authors' details

[1] Cancer Epidemiology Unit, Department of Medical Sciences, University of Turin and CPO Piemonte, Turin, Italy;

Introduction Several studies have reported an increased risk of wheezing in children of mothers who used paracetamol during pregnancy[1]–[4]. However, the lack of complete adjustment for confounding factors has been suggested as a possible explanation for the observed association[5]. Several factors can be associated with paracetamol use during pregnancy and, at the same time, may be related to wheezing in children. Traditional approaches for the estimation of the effect of an exposure on an outcome may be biased under model misspecification. For example, multivariable regression requires the correct specification of a parametric outcome model. Propensity score methods require the correct specification of both the exposure and, in some cases, the outcome model[6][7]. Targeted maximum likelihood estimation (TMLE) is a semiparametric doubly-robust method that allows to obtain an unbiased estimator of the effect if either the exposure or the outcome model is correctly specified, but not necessarily both[8]–[11]. Furthermore, TMLE allows the integration of machine learning methods in the estimation process. Machine learning methods may be useful to alleviate the bias due to statistical model misspecification as they avoid the assumption that the data-generating process has a specific probability distribution, and are generally flexible: they can handle interactions and nonlinearities, exploring a wide range of possible relationships[12]. In particular, TMLE can rely on the SuperLearner, an ensemble learning algorithm that combines multiple machine learning methods to create a robust and accurate prediction model[13][14].

Objectives In this paper, we apply TMLE in a birth cohort study, in a framework with a binary exposure and a binary outcome. The aim is the estimation of the effect of maternal paracetamol use during the first trimester of pregnancy on the occurrence of wheezing symptoms in the first 18 months of child' life in the NINFEA birth cohort.

Methods We applied three different implementations of TMLE: (i) parametric TMLE, using logistic regression for the exposure and the outcome model, (ii) nonparametric TMLE, with the use of default SuperLearner library, (iii) nonparametric TMLE, with the use of a user-defined SuperLearner library. For the purpose of comparison, we also applied traditional statistical approaches: (i) multivariable regression, (ii) propensity score regression adjustment, (iii) inverse probability of treatment weighting. We evaluated the effect on the risk ratio scale, including an increasing set of covariates. Covariate Set 1 includes socioeconomic baseline variables and maternal characteristics and health conditions (asthma and maternal allergic rhinitis) potentially associated with infant wheezing and possibly associated with paracetamol use during pregnancy. Covariate Set 2 additionally includes maternal infections (or factors associated with maternal infections) in the first trimester of pregnancy. Covariate Set 3 includes also pre-pregnancy maternal mental health, namely anxiety and depression. Covariate Set 4 further includes post-exposure potential confounders that were considered in previous studies, excluding potential mediators, namely factors that may be in the causal pathway between the exposure and the outcome. All the models were adjusted for these 4 different sets of covariates. We additionally evaluated a minimally adjusted estimate, which accounts for the region of residence only that we assumed to be the crude association.

Results We found that before adjustment, children exposed to paracetamol during pregnancy had a 26% increased risk of wheezing compared with those unexposed to paracetamol (RR 1.26; 95% CI (1.12;1.43)). After adjustment for gradually increasing set of covariates, the association between paracetamol and wheezing moves toward the null, and all statistical methods produce similar results. However, all the fully adjusted models, using both traditional and the 3 TMLE methods, suggest a positive, but weak, association (Table 1).

N=3934	Multivariable regression	Propensity score regression adjustment	IPW	TMLE		
	Poisson regression link=log			without SuperLearner	with default SuperLearner library	with user-defined SuperLearner library
Minimally adjusted	1.26 (1.12;1.43)	-	-	-	-	-
Set 1	1.16 (1.02;1.31)	1.16 (1.02;1.32)	1.16 (1.02;1.32)	1.16 (1.03;1.32)	1.17 (1.04;1.33)	1.18 (1.03;1.33)
Set 2	1.12 (0.98;1.27)	1.12 (0.99;1.28)	1.12 (0.97;1.28)	1.12 (0.98; 1.29)	1.13 (0.99;1.28)	1.13 (0.99;1.28)
Set 3	1.11 (0.98;1.26)	1.11 (0.98;1.27)	1.11 (0.96;1.27)	1.12 (0.97;1.29)	1.12 (0.99;1.28)	1.12 (0.99;1.27)
Set 4	1.11 (0.98;1.26)	1.12 (0.98;1.27)	1.11 (0.96;1.27)	1.12 (0.97;1.29)	1.11 (0.98;1.27)	1.12 (0.99;1.26)

Table 1: Association of maternal paracetamol use during the first trimester of pregnancy and risk of child wheezing at 18 months of age

Conclusion The association between maternal paracetamol use during pregnancy and infant wheezing is largely explained by confounding. However, a weak positive association remains. The use of TMLE paired with the SuperLearner can help to reduce concern about the bias due to statistical model misspecification. Residual confounding, potentially unmeasured, may explain the residual association, rather than any shortcomings in the statistical specification of the relationship between exposure, outcome, and confounders.

References

[1] S. O. Shaheen *et al.*, "Paracetamol use in pregnancy and wheezing in early childhood," *Thorax*, vol. 57, no. 11, pp. 958–963, Nov. 2002, doi: 10.1136/THORAX.57.11.958.

[2] L. Garcia-Marcos, M. Sanchez-Solis, V. Perez-Fernandez, M. D. Pastor-Vivero, P. Mondejar-Lopez, and J. Valverde-Molina, "Is the Effect of Prenatal Paracetamol Exposure on Wheezing in Preschool Children Modified by Asthma in the Mother?," *Int. Arch. Allergy Immunol.*, vol. 149, no. 1, pp. 33–37, Apr. 2009, doi: 10.1159/000176304.

[3] C. Vedel, D. S. Jørgensen, D. M. Kristensen, O. B. Petersen, and G. Greisen, "The use of paracetamol during pregnancy: A qualitative study and possible strategies for a clinical trial," *PLoS One*, vol. 17, no. 9, Sep. 2022, doi: 10.1371/JOURNAL.PONE.0271537.

[4] R. Patel, K. Sushko, J. van den Anker, and S. Samiee-zafarghandy, "Long-Term Safety of

Prenatal and Neonatal Exposure to Paracetamol: A Systematic Review," *International Journal of Environmental Research and Public Health*, vol. 19, no. 4. MDPI, Feb. 01, 2022, doi: 10.3390/ijerph19042128.

[5] E. Migliore *et al.*, "Prenatal Paracetamol Exposure and Wheezing in Childhood: Causation or Confounding?," 2015, doi: 10.1371/journal.pone.0135775.

[6] P. C. Austin, "An introduction to propensity score methods for reducing the effects of confounding in observational studies," *Multivariate Behav. Res.*, vol. 46, no. 3, pp. 399–424, May 2011, doi: 10.1080/00273171.2011.568786.

[7] M. A. Luque-Fernandez, M. Schomaker, B. Rachet, and M. E. Schnitzer, "Targeted maximum likelihood estimation for a binary treatment: A tutorial," *Stat. Med.*, vol. 37, no. 16, pp. 2530–2546, Jul. 2018, doi: 10.1002/SIM.7628.

[8] U. C. Berkeley, M. J. Van Der Laan, and D. Rubin, "Targeted Maximum Likelihood Learning." [Online]. Available: http://biostats.bepress.com/ucbbiostat/paper213.

[9] M. van der Laan and S. Gruber, "Targeted Minimum Loss Based Estimation of an Intervention Specific Mean Outcome," *U.C. Berkeley Div. Biostat. Work. Pap. Ser.*, Aug. 2011, Accessed: Dec. 20, 2022. [Online]. Available: https://biostats.bepress.com/ucbbiostat/paper290.

[10] S. Gruber and M. van der Laan, "Targeted Maximum Likelihood Estimation: A Gentle Introduction," *U.C. Berkeley Div. Biostat. Work. Pap. Ser.*, Aug. 2009, Accessed: Nov. 09, 2022. [Online]. Available: https://biostats.bepress.com/ucbbiostat/paper252.

[11] M. A. Luque-Fernandez, A. Belot, L. Valeri, G. Cerulli, C. Maringe, and B. Rachet, "Practice of Epidemiology Data-Adaptive Estimation for Double-Robust Methods in Population-Based Cancer Epidemiology: Risk Differences for Lung Cancer Mortality by Emergency Presentation," vol. 187, no. 4, doi: 10.1093/aje/kwx317.

[12] Q. Bi, K. E. Goodman, J. Kaminsky, and J. Lessler, "Practice of Epidemiology What is Machine Learning? A Primer for the Epidemiologist," doi: 10.1093/aje/kwz189.

[13] X. Tan, S. Yang, W. Ye, D. E. Faries, I. Lipkovich, and Z. Kadziola, "When Doubly Robust Methods Meet Machine Learning for Estimating Treatment Effects from Real-World Data: A Comparative Study," 2022.

[14] M. J. Van Der Laan, E. C. Polley, and A. E. Hubbard, "Super learner," *Stat. Appl. Genet. Mol. Biol.*, vol. 6, no. 1, Sep. 2007, doi: 10.2202/1544-6115.1309.