Observational Studies 4 (2018) 193-216 Submitted 12/17; Published 6/18

Comparing logistic and log-binomial models for causal mediation analyses of binary mediators and rare binary outcomes: evidence to support cross-checking of mediation results in practice

Mariia Samoilenko **mariia.samoilenko@uqam.ca**
Department of Mathematics, Université du Québec à Montréal
Research Center, Hôpital du Sacré-Coeur de Montréal
Montréal, Québec, Canada

Lucie Blais **lucie.blais@umontreal.ca**
Faculty of Pharmacy, Université de Montréal, Montréal
Research Center, Hôpital du Sacré-Coeur de Montréal, Montréal
Centre de recherche clinique Étienne-Le Bel, Centre hospitalier universitaire de Sherbrooke, Sherbrooke
Québec, Canada

Geneviève Lefebvre **lefebvre.gen@uqam.ca**
Department of Mathematics, Université du Québec à Montréal
Faculty of Pharmacy, Université de Montréal
Montréal, Québec, Canada

Abstract

Background: In the binary outcome framework to causal mediation, closed-form expressions introduced by Valeri and VanderWeele for the natural direct and indirect effect odds ratios (ORs) are established from a logistic outcome model by invoking several approximations that hold under the rare-disease assumption. Such ORs are expected to be close to corresponding effects on the risk ratio (RR) scale based on a log-binomial outcome model, however new insight indicates that this is not always verified. The objective was to report on mediation results from these two models when the incidence of the outcome was <10%.

Methods: Standard (approximate) ORs and RRs were estimated using data on a cohort of asthmatic pregnant women from Québec (Canada) and their babies. Prematurity and low birthweight were the mediator and outcome variables, respectively, and two binary exposure variables were considered: treatment to inhaled corticosteroids and placental abruption. Exact closed-form effects expressed on the OR scale were also derived and estimated using a SAS code we provide. A study based on two simulation scenarios was subsequently devised to supplement on the substantive findings.

Results: Many approximate ORs and RRs estimated from our cohort analyses did not closely agree. Approximate ORs were systematically observed farther from RRs in comparison with exact ORs, possibly leading to different conclusions regarding the null hypothesis. Exact OR estimates were very close to RR estimates for exposure to inhaled corticosteroids, but less so for placental abruption. The approximate OR estimator was found to exhibit important bias and undercoverage in the simulation scenario which featured a strong mediator-outcome relationship.

Conclusions: Logistic and log-binomial outcome models can yield dissimilar binary-binary mediation effects even if the outcome incidence is small marginally. Large discrepancies between approximate ORs and RRs may indicate invalid inference for these ORs. Exact OR estimates can be obtained for validation or to replace RRs if the log-binomial model exhibits convergence problems.

Keywords: Causal Mediation, Binary Outcome and Mediator, Rare-Disease Assumption, Natural Direct and Indirect Effects, Odds Ratio, Risk Ratio, Logistic Model, Log-Binomial Model, Approximation

1. Introduction

The incidence of an outcome of interest is a factor to consider when selecting a model to perform a causal mediation analysis of a binary outcome. When the outcome is rare, the Valeri and VanderWeele (2013) regression-based counterfactual approach to mediation relies on a logistic model for the outcome given the mediator, exposure, and confounders. In this familiar mediation framework (Wang and Arah, 2015), the model-based expressions for the natural direct and indirect effects of the exposure are given on the odds ratio (OR) scale. For both binary and continuous mediators, the closed-form formula for each of these effects is established using a series of approximations that are assumed to hold under the so-called rare-disease assumption. In mediation applications and in epidemiology more generally, a practical threshold for qualifying an outcome (disease) as rare is an incidence of less than 10% (e.g., VanderWeele, 2016). Although restrictive on the outcome, this mediation regression approach based on a logistic outcome model is appealing in practice due to its apparent conceptual simplicity.

A number of mediation analysis approaches that do not impose constraints on the rareness of the outcome are also available to practitioners. For instance, the simulation-based approach introduced by Imai et al. (2010) allows for mediation analyses with non-rare binary outcomes and returns effect estimates on a risk difference scale. The g-computation approach of Daniel et al. (2011) and the imputation approach of Vansteelandt et al. (2012) also do not require the studied binary outcome to be rare. These mediation approaches can return estimates on the OR or the risk ratio (RR) scales, the latter scales being arguably more familiar to health sciences practitioners than a risk difference scale when dealing with binary outcomes. Lastly, the closed-form regression-based mediation framework can also use a log-binomial (a.k.a. log-linear) outcome model as a replacement to the logistic model when the outcome is not rare (Valeri and VanderWeele, 2013; Ananth and VanderWeele, 2011). As warned by Valeri and VanderWeele (2013), with a non-rare outcome the OR does not approximate the RR anymore, and the proposed natural direct and indirect effect OR estimators will be biased for the RR if a logistic regression is used to model the outcome. As opposed to the logistic, the log-binomial model provides natural direct and indirect effect estimates on the RR scale exactly. These two modelling options, as well as the other aforementioned approaches, have all been implemented in macros for mediation analyses (Valeri and VanderWeele, 2013; Daniel et al., 2011; Starkopf et al., 2017; Steen et al., 2017; Tingley et al., 2014) and are thus widely available for epidemiologists and data analysts.

In this paper, we report on mediation results obtained from the aforementioned regression-based mediation framework using logistic and log-binomial outcome models in the case of a rare outcome using maternal and perinatal data recorded in administrative databases for

pregnant women with asthma from the province of Québec (Canada). Natural direct and indirect OR and RR effects obtained respectively from these two types of models have been suggested to be close when the rare-disease assumption holds (VanderWeele and Vansteelandt, 2014). However, considering the binary case for both the outcome and the mediator, we have observed that currently implemented *approximate* natural direct and indirect effect ORs based on a logistic outcome model (Valeri and VanderWeele (2013) macro, CAUSALMED procedure in SAS 9.4M5 and later releases) can substantially differ from corresponding RRs obtained from a log-binomial outcome model *even if the incidence of the outcome is less than 10%*. This phenomenon, which, to our knowledge, had not yet been documented, was revealed to us on the basis of two data analysis scenarios involving prematurity (PTB) as a mediator and low birthweight (LBW) as the outcome. The binary exposure considered in our first scenario was treatment with inhaled corticosteroids during pregnancy while placental abruption was considered in the second scenario. Because *exact* closed-form formulas of natural effects expressed on the OR scale can be derived from the logistic model, we then decided to investigate whether approximate and exact OR estimators would behave similarly. As detailed subsequently, exact OR inference simply consists in working with the exact definitions for the conditional natural direct and indirect ORs, as opposed to working with simplified expressions from these in the approximate OR approach.

The paper is organized as follows. We first review the studied regression-based counterfactual approach to mediation analysis under binary outcome and mediator. More specifically, we present the natural direct and indirect approximate OR and RR formulas pertaining to modeling the outcome using either the logistic or the log-binomial model, respectively, while the mediator is taken to be modeled with a logistic regression. We then introduce the exact OR expressions for the natural direct and indirect effects based on the logistic outcome model. Descriptive statistics on the exposure, mediator and outcome variables from our cohort are given, followed by the mediation results for the association between inhaled corticosteroids and LBW and placental abruption and LBW obtained from both models. Results from a simulation study based on two data-generation mechanisms are then presented to supplement on the substantive findings.

2. Methods

2.1 Natural direct and indirect effects

Natural direct and indirect effects are described in terms of so-called nested counterfactuals, $Y(a, M(a^*))$, which represent the outcome that would have been observed if exposure A were set to a and mediator M to the value it would have taken if A were set to a^*, where a^* indicates a reference or baseline value of A (Robins and Greenland, 1992; Pearl, 2001) At the unit level, the comparison between $Y(a, M(a^*))$ and $Y(a^*, M(a^*))$ entails the natural direct effect of changing exposure a to a^*, while the contrast between $Y(a, M(a))$ and $Y(a, M(a^*))$ represents the natural indirect effect. Such comparisons can be made on subpopulations conditional on covariates $\boldsymbol{C}$ (that is, in terms of average $E[Y(a, M(a^*))|\boldsymbol{C} = \boldsymbol{c}]$) and reported, for instance, using differences or using odds or risk ratios if the outcome is binary.

2.2 Standard binary-binary closed-form regression-based counterfactual approach to mediation

With a binary outcome Y and binary mediator M, the Valeri and VanderWeele (2013) counterfactual approach to mediation based on the logistic model assumes the following regression equations for the data:

$$\begin{aligned} logit\{P(Y=1|A=a,M=m,\boldsymbol{C}=\boldsymbol{c})\} &= \theta_0+\theta_1 a+\theta_2 m+\theta_3 am+\boldsymbol{\theta}_4^{'}\boldsymbol{c}, \quad (1)\\ logit\{P(M=1|A=a,\boldsymbol{C}=\boldsymbol{c})\} &= \beta_0+\beta_1 a+\boldsymbol{\beta}_2^{'}\boldsymbol{c}, \quad (2)\end{aligned}$$

where A is the exposure variable (either binary or continuous) and $\boldsymbol{C}$ are adjustment covariates. The set of covariates must be selected carefully for the identification of mediation effects and must be such that the following three assumptions hold: (i) no unmeasured treatment-outcome confounding; (ii) no unmeasured mediator-outcome confounding; and (iii) no unmeasured treatment-mediator confounding. A fourth assumption also requires that there be no mediator-outcome confounders affected by treatment. In addition to these confounding assumptions, the no-interference, consistency, and composition assumptions are required for identification. All these assumptions above ensure that nested potential outcome expectations can be calculated using the mediation formula $E[Y(a,M(a^*))|\boldsymbol{C}=\boldsymbol{c}]=\sum_m E[Y|A=a,M=m,\boldsymbol{C}=\boldsymbol{c}]P(M=m|A=a^*,\boldsymbol{C}=\boldsymbol{c})$. We refer the reader to VanderWeele and Vansteelandt (2009) and VanderWeele (2015) for further details on causal mediation assumptions.

If regression models (1) and (2) are correctly specified, the aforementioned identification assumptions hold, and the outcome Y is rare, then the (conditional) natural direct (NDE) and indirect (NIE) effects on the OR scale, comparing binary exposure level 1 to 0, can be approximated as (Valeri and VanderWeele, 2013; VanderWeele, 2015) :

$$\begin{aligned} OR_{NDE}^{app}(\boldsymbol{c}) &= \frac{exp(\theta_1)(1+exp(\theta_2+\theta_3+\beta_0+\boldsymbol{\beta}_2^{'}\boldsymbol{c}))}{1+exp(\theta_2+\beta_0+\boldsymbol{\beta}_2^{'}\boldsymbol{c})}, \quad (3)\\ OR_{NIE}^{app}(\boldsymbol{c}) &= \frac{(1+exp(\beta_0+\boldsymbol{\beta}_2^{'}\boldsymbol{c}))(1+exp(\theta_2+\theta_3+\beta_0+\beta_1+\boldsymbol{\beta}_2^{'}\boldsymbol{c}))}{(1+exp(\beta_0+\beta_1+\boldsymbol{\beta}_2^{'}\boldsymbol{c}))(1+exp(\theta_2+\theta_3+\beta_0+\boldsymbol{\beta}_2^{'}\boldsymbol{c})}. \quad (4)\end{aligned}$$

The natural direct effect (3) can be interpreted as the comparison of the odds of the outcome Y if each unit of the subpopulation with $\boldsymbol{C}=\boldsymbol{c}$ had been exposed, but the mediator had been fixed to what it would have been under no exposure, to the odds of the outcome Y if each unit had been unexposed and the mediator had been fixed at its natural level under no exposure. The natural indirect effect (4) can be interpreted, conditionally on $\boldsymbol{C}=\boldsymbol{c}$, as the comparison of the odds of the outcome Y if each unit of the subpopulation had been exposed, and the mediator had been fixed to what it would have been under exposure, to the odds of the outcome Y if each unit had been exposed, but the mediator had been fixed at its natural level.

Point estimates for $OR_{NDE}^{app}(\boldsymbol{c})$ and $OR_{NIE}^{app}(\boldsymbol{c})$ are obtained by replacing the parameters in (3-4) by their estimated values. Confidence intervals for these effects can be obtained by the delta method or the bootstrap (Casella and Berger, 2002; Chernick, 2011) .

Under a log-binomial model for the outcome, the studied regression-based approach instead assumes that

$$log\{P(Y=1|A=a,M=m,\boldsymbol{C}=\boldsymbol{c})\}=\theta_0^*+\theta_1^* a+\theta_2^* m+\theta_3^* am+\boldsymbol{\theta}_4^{*'}\boldsymbol{c}, \quad (5)$$

while the mediator model remains as in (2). The NDE and NIE on the RR scale are then expressed as:

$$RR_{NDE}(\boldsymbol{c}) = \frac{exp(\theta_1^*)(1+exp(\theta_2^*+\theta_3^*+\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c}))}{1+exp(\theta_2^*+\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})}, \quad (6)$$

$$RR_{NIE}(\boldsymbol{c}) = \frac{(1+exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c}))(1+exp(\theta_2^*+\theta_3^*+\beta_0+\beta_1+\boldsymbol{\beta}_2'\boldsymbol{c}))}{(1+exp(\beta_0+\beta_1+\boldsymbol{\beta}_2\prime\boldsymbol{c}))(1+exp(\theta_2^*+\theta_3^*+\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c}))}. \quad (7)$$

The interpretation of $RR_{NDE}(\boldsymbol{c})$ and $RR_{NIE}(\boldsymbol{c})$ is similar to that of $OR_{NDE}^{app}(\boldsymbol{c})$ and $OR_{NIE}^{app}(\boldsymbol{c})$ excepts that it compares risks instead of odds. The estimation of $RR_{NDE}(\boldsymbol{c})$ and $RR_{NIE}(\boldsymbol{c})$ also proceeds similarly.

2.3 Exact closed-form mediation effect formulas for binary-binary logistic models

One must appreciate that expressions (3) and (4) for the NDE and NIE are approximate formulas established by invoking the rare-disease assumption multiple times. When comparing (3) with (6) and (4) with (7), we remark that each pair possesses the same analytical form albeit different parameter values (i.e., $\boldsymbol{\theta}$ vs $\boldsymbol{\theta}^*$), leading to simplified interpretation and statistical inference for the ORs. Because the estimation of mediator regression coefficients $\boldsymbol{\beta}$ does not depend on the choice of outcome model, these ORs would thus be interpretable as RRs when all outcome regression coefficients in (3) and (4) are close to corresponding coefficients in (6) and (7), respectively. Nonetheless, it is straightforward to show that the NDE and NIE on the OR scale can be formulated exactly as:

$$OR_{NDE}(\boldsymbol{c}) = \frac{P(Y(1,M(0))=1|\boldsymbol{C}=\boldsymbol{c})/(1-P(Y(1,M(0))=1|\boldsymbol{C}=\boldsymbol{c}))}{P(Y(0,M(0))=1|\boldsymbol{C}=\boldsymbol{c})/(1-P(Y(0,M(0))=1|\boldsymbol{C}=\boldsymbol{c}))}, \quad (8)$$

$$OR_{NIE}(\boldsymbol{c}) = \frac{P(Y(1,M(1))=1|\boldsymbol{C}=\boldsymbol{c})/(1-P(Y(1,M(1))=1|\boldsymbol{C}=\boldsymbol{c}))}{P(Y(1,M(0))=1|\boldsymbol{C}=\boldsymbol{c})/(1-P(Y(1,M(0))=1|\boldsymbol{C}=\boldsymbol{c}))}, \quad (9)$$

where

$$\begin{aligned} P(Y(1,M(1))=1|\boldsymbol{C}=\boldsymbol{c}) &= P(Y=1|A=1,M=1,\boldsymbol{C}=\boldsymbol{c})\cdot P(M=1|A=1,\boldsymbol{C}=\boldsymbol{c}) \\ &\quad +P(Y=1|A=1,M=0,\boldsymbol{C}=\boldsymbol{c})\cdot P(M=0|A=1,\boldsymbol{C}=\boldsymbol{c}) \\ &= \frac{exp(\theta_0+\theta_1+\theta_2+\theta_3+\boldsymbol{\theta}_4'\boldsymbol{c})}{1+exp(\theta_0+\theta_1+\theta_2+\theta_3+\boldsymbol{\theta}_4'\boldsymbol{c})}\cdot\frac{exp(\beta_0+\beta_1+\boldsymbol{\beta}_2'\boldsymbol{c})}{1+exp(\beta_0+\beta_1+\boldsymbol{\beta}_2'\boldsymbol{c})} \\ &\quad +\frac{exp(\theta_0+\theta_1+\boldsymbol{\theta}_4'\boldsymbol{c})}{1+exp(\theta_0+\theta_1+\boldsymbol{\theta}_4'\boldsymbol{c})}\cdot\frac{1}{1+exp(\beta_0+\beta_1+\boldsymbol{\beta}_2'\boldsymbol{c})}, \end{aligned} \quad (10)$$

$$
\begin{aligned}
P(Y(1,M(0))=1|\boldsymbol{C}=\boldsymbol{c}) &= P(Y=1|A=1,M=1,\boldsymbol{C}=\boldsymbol{c})\cdot P(M=1|A=0,\boldsymbol{C}=\boldsymbol{c}) \\
&\quad +P(Y=1|A=1,M=0,\boldsymbol{C}=\boldsymbol{c})\cdot P(M=0|A=0,\boldsymbol{C}=\boldsymbol{c}) \\
&= \frac{exp(\theta_0+\theta_1+\theta_2+\theta_3+\boldsymbol{\theta}_4'\boldsymbol{c})}{1+exp(\theta_0+\theta_1+\theta_2+\theta_3+\boldsymbol{\theta}_4'\boldsymbol{c})}\cdot\frac{exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})}{1+exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})} \\
&\quad +\frac{exp(\theta_0+\theta_1+\boldsymbol{\theta}_4'\boldsymbol{c})}{1+exp(\theta_0+\theta_1+\boldsymbol{\theta}_4'\boldsymbol{c})}\cdot\frac{1}{1+exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})}, \qquad (11)
\end{aligned}
$$

$$
\begin{aligned}
P(Y(0,M(0))=1|\boldsymbol{C}=\boldsymbol{c}) &= P(Y=1|A=0,M=1,\boldsymbol{C}=\boldsymbol{c})\cdot P(M=1|A=0,\boldsymbol{C}=\boldsymbol{c}) \\
&\quad +P(Y=1|A=0,M=0,\boldsymbol{C}=\boldsymbol{c})\cdot P(M=0|A=0,\boldsymbol{C}=\boldsymbol{c}) \\
&= \frac{exp(\theta_0+\theta_2+\boldsymbol{\theta}_4'\boldsymbol{c})}{1+exp(\theta_0+\theta_2+\boldsymbol{\theta}_4'\boldsymbol{c})}\cdot\frac{exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})}{1+exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})} \\
&\quad +\frac{exp(\theta_0+\boldsymbol{\theta}_4'\boldsymbol{c})}{1+exp(\theta_0+\boldsymbol{\theta}_4'\boldsymbol{c})}\cdot\frac{1}{1+exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})}. \qquad (12)
\end{aligned}
$$

We refer the reader to the Appendix for further details on how one can start from the exact ORs (8-9) and end with the approximate ORs (3-4). When comparing approximate ORs (3-4) with exact ORs (8-9) combined with (10-12), one immediately appreciates the lesser complexity of the expressions pertaining to the approximate ORs, which facilitates statistical inference. We conjecture that an important incentive for the wide adoption of the approximate OR approach in practice was the availability of standard error formulas based on the delta method, which is lacking for the exact OR inference thus far.

2.4 Data and variables

The data used in this paper come from two administrative databases in Québec: the Régie de l'assurance maladie du Québec (RAMQ) and the Maintenance et exploitation des données pour l'étude de la clientèle hospitalière (MED-ECHO) databases. Our cohort comprises pregnancies from asthmatic women who delivered between 1998 and 2008 and is a subset of pregnancies used in the study by Samoilenko et al. (2016); inclusion and exclusion criteria for the larger cohort can be found in this reference. Our analyses used the most recent pregnancy per woman during the follow-up period, for a total of n=6197 pregnancies. LBW and PTB measures were available for all pregnancies, as well as exposure information on use of inhaled corticosteroids during pregnancy (yes/no) and occurrence of placental abruption (yes/no). Both unadjusted (crude) and adjusted analyses were planned to verify whether our findings would be affected by the inclusion of covariates in the outcome and mediator models; as such, two sets of covariates which differed in size (4 versus 27) were considered in (1) (or (5)) and (2) (see Table A1 in Appendix). Of note, in the inhaled corticosteroids exposure analysis, placental abruption was considered as a potential confounder for PTB

and LBW and was therefore included in corresponding large set analysis, and vice-versa for the placental abruption exposure analysis.

2.5 Statistical analyses

Descriptive statistics were obtained to describe marginal and conditional probabilities for the outcome (LBW), mediator (PTB) and both exposures (inhaled corticosteroids and placental abruption) in our cohort.

Using the macro by Valeri and VanderWeele (2013), we obtained the unadjusted natural direct and indirect OR effect estimates of each exposure on LBW, with PTB as the mediator variable, based on standard expressions (3) and (4), respectively. The total effect of each exposure was also obtained and corresponded to the multiplication of the NDE and NIE (that is, TOTAL=NDE×NIE). Confidence intervals (at 95%) were calculated using the delta method and the bootstrap with percentile approach using 1000 replications. Similarly, we obtained the unadjusted RR results using the Valeri and VanderWeele (2013) macro. This was done by using the log-binomial (log-linear) option for the outcome model; note that this macro allows only for a logistic model for the binary mediator. Bootstrap intervals for the RRs were obtained with the same random seeds as for the ORs. Exact inference for the ORs was based on the formulas presented in (8-9), and the bootstrap method with 1000 replications was also used to construct 95% confidence intervals. The whole analysis was then repeated for each set of adjustment covariates, where results are reported conditionally on the mean level of the covariates $\boldsymbol{C}$. For these adjusted analyses, we manually added an option to the log-binomial outcome model statement (intercept=-4). This option is proposed in the SAS documentation to help ensure probabilities between 0 and 1 for the reference level when the log-binomial estimation procedure is started (Fang, 2011).

All exact OR results were generated using a contributed SAS code inspired by the Valeri and VanderWeele (2013) macro. In our code, unlike in this macro, the bootstrap datasets are generated using the SAS procedure PROC SURVEYSELECT which has been described as an efficient way to implement bootstrap computations (Cassell, 2007). Our code is made available in the paper's Supplemental Material and is illustrated using a large synthetic dataset generated according to each of the two simulation scenarios considered herein. The exact OR results obtained on our cohort were also compared to the OR results returned by the imputation mediation approach implemented in the `R` package `medflex` (Steen et al., 2017). For this approach, we used the outcome model specification (1) to create an expanded imputed dataset on which the following conditional mean model for the nested counterfactuals $Y(a, M(a^*))$, for $a, a^* = 0, 1$, was fitted:

$$logit\{P(Y(a, M(a^*)) = 1|\boldsymbol{C} = \boldsymbol{c})\} = \gamma_0 + \gamma_1 a + \gamma_2 a^* + \gamma_3 a a^* + \boldsymbol{\gamma}_4^{'}\boldsymbol{c}. \quad (13)$$

Our combined choice of outcome model (1) and natural effect model (13) followed the recommendation of Vansteelandt et al. (2012) in that the former should at least reflect the structure of the latter, with a^* substituted by M. Results using a log link function instead of a logit in (13) were also obtained to assess the correspondence between this alternative imputation approach and the log-binomial approach yielding equations (6-7).

Approval from the Commission d'accès à l'information du Québec was obtained prior to requesting and linking the information from the MED-ECHO and RAMQ databases. These analyses were approved by the ethics committee of the Hôpital du Sacré-Coeur de Montréal.

2.6 Design of simulation study

For completeness, we compared the performance of the approximate and exact OR approaches in two simulated scenarios that both featured a rare outcome marginally. The data-generation mechanism in each scenario is summarized in Table 1 and yields a marginal probability of the outcome approximatively equal to 9.5% in each of them. Scenario 1 corresponds to a scenario in which the conditional probability of the outcome is relatively small in each of the four strata corresponding to combinations of levels of the exposure and mediator. Scenario 2 is such that the conditional probability of the outcome is large for one stratum of the mediator, as could be found in our real data examples. For each simulation scenario, 1000 samples of size n=2000 or 10,000 were nonparametrically generated using sequential Bernoulli sampling for variables A, M, and Y. Approximate and exact closed-form estimates of natural effects expressed on the OR scale were obtained on each sample as in the substantive examples.

For each simulation scenario, the true mediation RRs and ORs were calculated from Table 1. The bias, standard deviation, root mean square error (RMSE), and coverage probability of the approximate and exact OR estimators were then estimated based on the 1000 samples generated; the true mediation RRs as well as the true mediation ORs were used as reference values.

3. Results

3.1 Results from cohort analyses

The marginal and conditional probabilities for the outcome (LBW), mediator (PTB) and both exposures (inhaled to corticosteroids and placental abruption) are presented in Table 2. LBW and PTB were observed in 7.54% and 9.29% of the pregnancies constituting our cohort, respectively, both smaller than the 10% threshold for rare outcomes. We remark, however, that while LBW is rare marginally, it is very common among PTB infants; this is the case also when conditioning on each level of both exposures. In particular, occurrence of LBW was seen to exceed 70% among PTB pregnancies with placental abruption. Unlike for pregnancies exposed to inhaled corticosteroids, PTB was not rare among pregnancies with placental abruption (32.27%).

Table 3 presents, for each exposure, outcome regression coefficients based on the logistic and log-binomial models for the crude and small set adjusted analyses (results for the largest set of covariables are not presented due to convergence problems in the log-binomial model). More precisely, regression coefficients pertaining to exposure (θ_1, θ_1^*), mediator (θ_2, θ_2^*), and exposure-mediator interaction (θ_3, θ_3^*) appearing in expressions (3-4) and (6-7) are shown in these tables. For exposure to inhaled corticosteroids, we found that the logistic and log-binomial regression coefficients associated with the mediator variable did not closely agree, both in the crude and small set adjusted analyses. For placental abruption, we also found notable discrepancies between the two models, with, in particular, large relative differences for the exposure-mediator interaction regression coefficients.

Tables 4-5 present the mediation results for our analyses. These tables reveal that many approximate OR and RR estimates did not closely agree, as could be expected from results shown in Table 3. Moreover, in all analyses where RR results were available, the approxi-

mate ORs were systematically observed to be farther from corresponding RRs in comparison with exact ORs. Exact ORs were very close to RR estimates for exposure to inhaled corticosteroids, but less so for placental abruption. Exact ORs were also systematically farther away from the null than RRs, except for two estimates in the smallest adjusted analysis for exposure to inhaled corticosteroids. In these two cases, exact ORs and RRs were nevertheless very close to each other.

For exposure to inhaled corticosteroids more specifically, the approximate OR NDE was found smaller than the corresponding RR and exact OR in both the unadjusted and small set adjusted analyses. For instance, the unadjusted approximate OR NDE point estimate was 0.843 while the same NDE on the RR and exact OR scales were 0.999 and 0.998, respectively. Moreover, in the fully adjusted analysis, the approximate OR for NDE was 0.682 while the corresponding exact OR was 0.862 (RR result not available). For this adjusted approximate OR, we rejected the null hypothesis of no effect (95% CI: 0.491, 0.947) while we did not for the exact OR (95% CI: 0.704, 1.055). For all reported results for exposure to inhaled corticosteroids, the length of the confidence intervals associated with the approximate ORs was also larger than those associated with the RRs and exact ORs.

For placental abruption, we observed that both approximate and exact ORs did not approximate the RRs well. The discrepancy between the approximate ORs and RRs was however larger than the one observed between the exact ORs and RRs. For example, in the small set adjusted analysis, the relative error for the approximate OR NIE point estimate for interpretation as measure of RR was 38.4% ((3.026-2.186)/2.186) while it was 23.3% for the exact OR NIE ((2.696-2.186)/2.186). Bootstrap confidence intervals for the adjusted approximate ORs exhibited excessively large upper bounds due to convergence problems in a non-negligible number of bootstrap datasets. In general, some numerical instability was observed for all measures across bootstrap datasets.

The results for the imputation approach based on models (1) and (13) are presented in the last two columns of Tables 4-5. The results obtained with this approach were overall similar to those obtained with the exact OR approach, but a larger discrepancy between results was observed for the large set adjusted analysis with placental abruption exposure. In general, some discrepancy in the adjusted analyses results was to be expected given that model specification (13) of the imputation approach implies constant OR mediation effects across covariate levels, unlike the regression-based ones (for both approximate and exact OR closed-forms). For the unadjusted analyses, results from the imputation approach with the log link function were very similar to those obtained with the regression-based approach with a log-binomial outcome model (results not shown). The imputation approach bootstrap CIs (with log link) were not returned by `medflex` for the small set adjusted analysis for inhaled corticosteroids exposure due to numerical problems. Moreover, contrary to the log-binomial approach using Valeri and VanderWeele (2013) macro with the intercept initial value = -4 in the outcome model, `medflex` did not return results (point estimates and CIs) for the small set adjusted analysis for the placental abruption exposure. For both exposures, no results could be obtained from `medflex` for the large set adjusted analysis, as was also the case with the log-binomial approach.

To concentrate on the impact of the rare-disease assumption, we have avoided so far discussing aforementioned confounding assumptions within the context of our analyses. Of course, it is unlikely that the unadjusted and small set adjusted analyses are sufficient

to fulfill confounding assumptions (i)-(iii), thus hindering the causal interpretation of corresponding estimates. Moreover, the assumption of no mediator-outcome confounders affected by treatment is also highly questionable. For instance, use of inhaled corticosteroids may influence uptake of other asthma medications that could then influence both PTB and LBW. The quality of a causal mediation analysis is evidently connected to the respect of these assumptions. While one often has no or only limited control on these, it is otherwise for the rare-disease assumption discussed herein.

3.2 Results from simulation study

The simulation study results for both scenarios are presented in Tables 6-7, where each table corresponds to results obtained using either the true mediation RRs or the true mediation ORs as reference, respectively.

For each effect in Scenario 1, we observed that the bias of the approximate OR estimator was systematically slightly larger than the bias of the exact OR estimator, as interpreted as a RR (top part of Table 6). The standard deviation of the approximate OR estimator was also larger than that of the exact OR. In Scenario 1, the coverage probabilities associated with the natural direct and indirect effects obtained from the approximate OR approach were close to the 95% nominal value when n=2000, but decreased below 90% with n=10,000. In comparison, the natural direct and indirect effect coverage probabilities for the exact OR approach stayed closer to 95% at both sample sizes, indicating the increased adequacy of this approach to yield estimates interpretable as RRs. Similar comments regarding the comparative performances of the approximate and exact OR estimators could be done for Scenario 1 from Table 7, while globally noting smaller biases and better coverage probabilities when taking the true mediation ORs as reference.

The intrinsic features of the estimators were exacerbated in Scenario 2. Large biases were seen for the approximate OR estimator and poor coverage probabilities were obtained with either the true RRs or ORs as references. A striking result is the 0.4% coverage probability for the natural direct effect from this approach when n=10,000 (with the true direct effect RR as reference); in contrast, the exact OR approach yielded a coverage probability of 92.4%. Important to mention is the adequacy of the exact OR approach to return estimates interpretable as OR, with coverage probabilities near 95% when taking the true ORs as reference (see bottom part of Table 7).

4. Discussion

In this paper, we have provided evidence that extra consideration is warranted when estimating natural direct and indirect effects with binary mediators and outcomes. Approximate ORs and RRs returned by studied standard logistic and log-binomial regression-based mediation approaches did not closely agree in our cohort analyses, although it was expected otherwise since the incidence of the outcome was less than 10%. In these analyses, the degree of agreement between the exact ORs and the RR estimates varied according to the nature of exposure, but was better than the one observed between the approximate ORs and the RRs. We have also made the troubling observation that approximate OR inference can lead to different conclusions regarding the null hypothesis of absence of effect as opposed to inference based on the exact OR scale. This indicates that data analysts must be cau-

tious when applying mediation models in this context. Although regularly invoked in the causal mediation literature, the rare-disease assumption is rarely, if ever, explicitly defined. While the 10% marginal threshold is commonly used in epidemiology, herein it failed to provide valid grounds for the approximate ORs mediation approach. A more precise but more abstract definition of a rare outcome is an incidence which implies the approximate equivalence between the logit and the log of the probability of the occurrence of the outcome (e.g., VanderWeele and Vansteelandt, 2010; Valeri and VanderWeele, 2013). As seen in the Appendix, the approximate OR formulas (3-4) are derived from exact formulas (8-9) by repeatedly invoking this equivalence. What should be noted, however, is that this equivalence must hold conditionally on variables, including the mediator, throughout the derivation of equations (3-4). Indeed, the fact that the rare-disease assumption needs to hold in all strata formed by conditioning variables is already known, but perhaps underappreciated, in general applications based on the logistic regression model (Cummings, 2009).

One hypothesis we put forward to explain the results we obtained in our cohort analyses is the strong relationship between the mediator (PTB) and the outcome (LBW): while the probability of LBW is small in the cohort overall, it is large for babies that are PTB, thus violating the rare-disease assumption in one stratum of the mediator. Indeed, this is a likely explanation to the discrepancy observed between regression coefficients associated with the mediator in the logistic and log-binomial outcome models, coefficients which are involved in the approximate OR and RR mediation effect expressions (3-4) and (6-7). As seen in our analyses, disagreement between mediation effect measures can be further exacerbated in presence of significant exposure-mediator interaction or when conditioning on covariates. One studied simulation scenario (Scenario 2), which violated the rare-disease assumption in one stratum of the mediator, also led to results similar to those found on the basis of our cohort data (recall Tables 6-7). In that scenario, the approximate OR estimator was found highly biased and showing very poor coverage performance. The logical consequence of these results would be that standard mediation applications based on the logistic model should preclude considerations of mediators that are strongly associated with the outcome. With this in mind, it will thus be interesting to further characterize how the discrepancy between the approximate OR and RR (or exact OR) results may arise in practice.

In the meantime, we advise data analysts to systematically obtain mediation effects from both logistic and log-binomial outcome models even if the 10% threshold for the outcome is satisfied marginally. Using the SAS code we provide, exact OR estimates can also be obtained for validation purposes or to replace RRs if the log-binomial model exhibits convergence problems. Further developing exact OR inference will be important since convergence problems with the log-binomial model are notorious (e.g., Localio et al., 2007). Indeed, taking advantage of the inferential stability of the logistic model while ensuring proper interpretation of mediation results would seem important. Other mediation approaches that do not rely on the rare-disease assumption may also be considered for validation. In this work, our `medflex` implementation of the imputation approach by Vansteelandt et al. (2012) was found aligned with the proposed exact OR approach. While the goal here was not to provide extensive comparisons between these two approaches, it is nevertheless appropriate to mention that the proposed exact OR approach can be easily conceptualized with general exposure variables, similarly to the approximate OR approach (Valeri and VanderWeele, 2013), while the imputation approach is less amenable to continuous exposures (Vansteelandt et al.,

2012). Moreover, we agree with Starkopf et al. (2017) that, in practice, the choice of estimation method is often based on software preference; as such, our SAS macro can be useful for applied researchers for whom SAS is the first choice for conducting mediation analyses.

We have seen from our cohort analysis that the two approaches studied which directly targeted the natural direct and indirect effects on the RR scale exhibited numerical problems even when adjusting for a small number of covariates. These two approaches were characterized by the use of a log link function specification in their implementation (recall model (5) and model (13) with log instead of logit). Although RRs are widely accepted to be more interpretable than ORs, the results obtained herein exposed difficulties with current approaches that are conceived exact for the former scale. Additional practical problems can thus be faced for those who are not willing to use, for instance, the proposed exact OR approach or the imputation approach with a logit link function to obtain mediation results. This suggests that numerically stable estimation strategies for binary outcome should also be put forward for the estimation of mediation effects on the RR scale.

To conclude, this paper provides strong incentive for moving away from the exclusive consideration of standard (approximate) closed-form OR formulas in binary-binary causal mediation analyses. Our work underlined the opaque behavior of the approximate OR approach in settings where it was deemed reasonable to apply at first sight. While other data could lead to only negligible differences between the approximate OR approach and the other approaches investigated herein, our substantive and simulated results are concerning enough to support aforementioned systematic cross-checking of binary-binary mediation results. We believe that our proposal amounts to good statistical practice, for which robustness of statistical outputs is assessed and reported results validated. Finally, although we have only considered the binary-binary case in our study, it is reasonable to expect that similar cross-checking would need to be done when using a continuous mediator and a binary outcome. This will need to be verified in the future.

Key Messages

- The Valeri and VanderWeele logistic regression-based counterfactual approach to mediation for a binary outcome and a binary mediator, derived from rare-disease approximations, can fail to provide natural direct and indirect effect odds ratios (ORs) interpretable as relative risks even when the outcome has a small incidence (<10%).
- To ascertain results, data analysts using this approach should systematically compare mediation effects obtained from a logistic model for the outcome and those obtained when instead using a log-binomial model.
- Exact closed-form estimates of natural effects expressed on the OR scale, which do not invoke rare-disease approximations, can additionally be obtained for purposes of validation using the contributed SAS macro *%OR_exact_formulas*.

Acknowledgments

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and from the Fonds de recherche du Québec - Santé (FRQS).

Table 1: Data simulation: probabilities for the exposure, mediator and outcome in Scenario 1 and Scenario 2

Probability	**Scenario 1**	**Scenario 2**
$P(A = 1)$	0.50	0.15
$P(M = 1\|A = 0)$	0.10	0.10
$P(M = 1\|A = 1)$	0.15	0.15
$P(Y = 1\|A = 0, M = 0)$	0.07	0.05
$P(Y = 1\|A = 0, M = 1)$	0.10	0.40
$P(Y = 1\|A = 1, M = 0)$	0.10	0.05
$P(Y = 1\|A = 1, M = 1)$	0.21	0.70

Table 2: Marginal and conditional probabilities (%) for outcome (Y) low birthweight, mediator (M) prematurity, separately for exposures (A) to inhaled corticosteroids and placental abruption among asthmatic pregnant women from Québec, Canada, 1998-2008

	Exposure: ICS	**Exposure: Placental Abruption**
$P(Y = 1)$	7.54	7.54
$P(M = 1)$	9.29	9.29
$P(A = 1)$	55.85	3.55
$P(M = 1\|A = 0)$	9.65	8.45
$P(M = 1\|A = 1)$	9.01	32.27
$P(Y = 1\|M = 0)$	2.78	2.78
$P(Y = 1\|M = 1)$	53.99	53.99
$P(Y = 1\|A = 0, M = 0)$	2.39	2.67
$P(Y = 1\|A = 0, M = 1)$	57.58	51.68
$P(Y = 1\|A = 1, M = 0)$	3.08	6.71
$P(Y = 1\|A = 1, M = 1)$	50.96	70.42

Table 3: Outcome model regresion coefficients: logistic regression model vs log-binomial regression model

Unadjusted (crude) analyses				**Adjusted analyses: small set**			
Regression coefficient	Logistic model	Log-binomial model	Relative difference[d] (%)	Regression coefficient	Logistic model	Log-binomial model	Relative difference[d] (%)
	Exposure: ICS				***Exposure: ICS***		
A^a	0.26	0.26	2.75	A^a	0.28	0.26	5.75
M^b	4.02	3.18	26.20	M^b	4.09	3.20	27.76
$A * M^c$	-0.53	-0.38	40.32	$A * M^c$	-0.55	-0.42	31.19
	Exposure: PA				***Exposure: PA***		
A^a	0.97	0.92	4.66	A^a	1.00	0.94	6.29
M^b	3.66	2.96	23.62	M^b	3.73	2.96	25.85
$A * M^c$	-0.17	-0.61	-73.08	$A * M^c$	-0.25	-0.68	-63.97

Abbreviations: ICS, inhaled corticosteroids; PA, placental abruption

[a]: θ_1 in logistic model (1), θ_1^* in log-binomial model (5);

[b]: θ_2 in logistic model (1), θ_2^* in log-binomial model (5);

[c]: θ_3 in logistic model (1), θ_3^* in log-binomial model (5);

[d]: $(\theta_i - \theta_i^*)/\theta_i^*$, $i = 1, 2, 3$.

Table 4: Unadjusted and adjusted natural direct, indirect, and total effects of treatment to inhaled corticosteroids on low birthweight mediated by prematurity based on approximate ORs, RRs, and exact ORs[a]

Effect	**App. OR**	Delta 95% CI	Bootstrap 95% CI	**RR**	Delta 95% CI	Bootstrap 95% CI	**Exact OR**	Bootstrap 95% CI	**Relative difference (%)** App. OR vs. RR[b]	Ex. OR vs. RR[c]	App. OR vs. ex. OR[d]	**OR (I)**	Bootstrap 95% CI
Unadjusted (crude) analyses													
NDE	**0.84**	0.64, 1.11	0.64, 1.16	**1.00**	0.87, 1.15	0.87, 1.16	**1.00**	0.85, 1.18	-15.62	-0.10	-15.53	**1.00**	0.85, 1.17
NIE	**0.95**	0.85, 1.07	0.84, 1.07	**0.96**	0.88, 1.05	0.88, 1.05	**0.96**	0.87, 1.05	-1.14	-0.42	-0.73	**0.96**	0.86, 1.06
TE[e]	**0.80**	0.59, 1.08	0.61, 1.09	**0.96**	0.81, 1.14	0.81, 1.14	**0.96**	0.79, 1.16	-16.48	-0.31	-16.21	**0.96**	0.78, 1.17
Adjusted analyses: small set													
NDE	**0.84**	0.63, 1.11	0.62, 1.15	**0.98**	0.85, 1.12	0.86, 1.14	**1.00**	0.85, 1.17	-14.45	2.05	-16.16	**1.00**	0.85, 1.18
NIE	**0.94**	0.84, 1.06	0.83, 1.06	**0.95**	0.87, 1.05	0.87, 1.04	**0.95**	0.87, 1.05	-1.47	-0.52	-0.95	**0.95**	0.85, 1.05
TE[e]	**0.79**	0.58, 1.07	0.58, 1.08	**0.93**	0.78, 1.11	0.79, 1.11	**0.94**	0.77, 1.15	-15.57	1.40	-16.74	**0.95**	0.78, 1.17
Adjusted analyses: large set													
NDE	**0.68**	0.49, 0.95	0.49, 0.98	**n/a**	n/a	n/a	**0.86**	0.70, 1.06	n/a	n/a	-21.00	**0.87**	0.70, 1.06
NIE	**0.89**	0.78, 1.02	0.78, 1.02	**n/a**	n/a	n/a	**0.91**	0.81, 1.02	n/a	n/a	-1.66	**0.92**	0.80, 1.05
TE[e]	**0.61**	0.43, 0.87	0.40, 0.90	**n/a**	n/a	n/a	**0.78**	0.62, 0.99	n/a	n/a	-22.18	**0.79**	0.62, 1.02

Abbreviations: app., approximate; CI, confidence interval; ex., exact; I, imputation; NDE, natural direct effect; NIE, natural indirect effect; OR, odds ratio; RR, relative risk; TE, total effect.

[a]: The approximate ORs are obtained from the standard regression-based approach with a logistic outcome model, the RRs are obtained from the standard regression-based approach with a log-binomial outcome model, and the exact ORs are obtained from proposed approach based on a logistic outcome model. ORs obtained by an imputation approach are presented for validation.

[b]: (approximate OR - RR)/RR;

[c]: (exact OR - RR)/RR;

[d]: (approximate OR - exact OR)/exact OR;

[e]: TE=NDE·NIE.

Table 5: Unadjusted and adjusted natural direct, indirect, and total effects of placental abruption on low birthweight mediated by prematurity based on approximate ORs, RRs, and exact ORs[a]

Effect	**App. OR**	Delta 95% CI	Bootstrap 95% CI	**RR**	Delta 95% CI	Bootstrap 95% CI	**Exact OR**	Bootstrap 95% CI	**Relative difference (%)** App. OR vs. RR[b]	Ex. OR vs. RR[c]	App. OR vs. ex. OR[d]	**OR (I)**	Bootstrap 95% CI
Unadjusted (crude) analyses													
NDE	**2.31**	1.49, 3.59	1.46, 4.22	**1.78**	1.29, 2.45	1.25, 2.42	**1.88**	1.26, 2.63	30.18	6.02	22.78	**1.88**	1.28, 2.68
NIE	**3.06**	2.41, 3.89	2.38, 3.91	**2.26**	1.74, 2.92	1.76, 3.05	**2.73**	2.06, 3.80	35.70	20.89	12.25	**2.73**	2.07, 3.76
TE[e]	**7.08**	4.15, 12.08	4.05, 14.81	**4.01**	3.17, 5.07	3.11, 4.94	**5.13**	3.64, 7.03	76.70	28.14	37.90	**5.13**	3.68, 6.87
Adjusted analyses: small set													
NDE	**2.24**	1.44, 3.50	1.39, 27441	**1.75**	1.26, 2.44	1.22, 2.39	**1.88**	1.23, 2.63	28.30	7.60	19.23	**1.90**	1.29, 2.70
NIE	**3.03**	2.37, 3.86	2.30, 3.91	**2.19**	1.69, 2.83	1.72, 2.99	**2.70**	2.02, 3.86	38.43	23.33	12.24	**2.71**	2.06, 3.75
TE[e]	**6.79**	3.95, 11.66	3.64, 107378	**3.82**	3.03, 4.83	2.96, 4.67	**5.07**	3.51, 6.90	77.66	32.73	33.85	**5.13**	3.69, 6.92
Adjusted analyses: large set													
NDE	**1.94**	1.14, 3.31	0.95, 406831	**n/a**	n/a	n/a	**1.70**	0.84, 2.58	n/a	n/a	14.31	**1.69**	1.08, 2.64
NIE	**1.83**	1.38, 2.42	1.31, 2.55	**n/a**	n/a	n/a	**1.65**	1.04, 2.34	n/a	n/a	11.25	**1.75**	1.33, 2.36
TE[e]	**3.55**	1.90, 6.62	1.61, 931013	**n/a**	n/a	n/a	**2.79**	1.06, 4.18	n/a	n/a	27.09	**2.96**	1.91, 4.70

Abbreviations: app., approximate; CI, confidence interval; ex., exact; I, imputation; NDE, natural direct effect; NIE, natural indirect effect; OR, odds ratio; RR, relative risk; TE, total effect.

[a]: The approximate ORs are obtained from the standard regression-based approach with a logistic outcome model, the RRs are obtained from the standard regression-based approach with a log-binomial outcome model, and the exact ORs are obtained from proposed approach based on a logistic outcome model. ORs obtained by an imputation approach are presented for validation.

[b]: (approximate OR - RR)/RR;

[c]: (exact OR - RR)/RR;

[d]: (approximate OR - exact OR)/exact OR;

[e]: TE=NDE·NIE.

Table 6: Simulation results: exact OR versus approximate OR estimators with RR as reference measure

		Exact OR results					**Approximate OR results**				
Effect	**True RR**[a]	Mean	Bias	SD	RMSE	CP	Mean	Bias	SD	RMSE	CP
Scenario 1: n=2000											
NIE	**1.050**	1.056	0.007	0.026	0.027	0.959	1.062	0.012	0.029	0.032	0.955
NDE	**1.521**	1.600	0.080	0.261	0.273	0.941	1.620	0.099	0.264	0.282	0.939
TE[b]	**1.596**	1.690	0.094	0.275	0.291	0.938	1.719	0.123	0.283	0.308	0.925
Scenario 1: n=10,000											
NIE	**1.050**	1.056	0.007	0.011	0.013	0.924	1.061	0.012	0.013	0.018	0.866
NDE	**1.521**	1.582	0.062	0.105	0.122	0.936	1.602	0.082	0.106	0.134	0.898
TE[b]	**1.596**	1.671	0.075	0.110	0.133	0.914	1.700	0.105	0.114	0.154	0.871
Scenario 2: n=2000											
NIE	**1.283**	1.346	0.063	0.173	0.184	0.933	1.418	0.136	0.198	0.240	0.891
NDE	**1.353**	1.411	0.059	0.229	0.236	0.947	2.650	1.297	0.865	1.559	0.384
TE[b]	**1.735**	1.893	0.157	0.348	0.382	0.932	3.763	2.028	1.339	2.430	0.402
Scenario 2: n=10,000											
NIE	**1.283**	1.334	0.051	0.075	0.091	0.903	1.407	0.125	0.089	0.153	0.676
NDE	**1.353**	1.396	0.043	0.101	0.110	0.924	2.497	1.144	0.341	1.194	0.004
TE[b]	**1.735**	1.861	0.126	0.163	0.206	0.866	3.519	1.784	0.554	1.868	0.003

Abbreviations: CP, coverage probability; NDE, natural direct effect; NIE, natural indirect effect; OR, odds ratio; RMSE, root mean square error; RR, relative risk; SD, standard deviation; TE, total effect.

[a]: True RRs are calculated from Table 1. For example, for Scenario 1:
true RR NIE = (0.21·0.15+0.10·(1-0.15))/(0.21·0.10+0.10·(1-0.10))=1.050,
true RR NDE = (0.21·0.10+0.10·(1-0.10))/(0.10·0.10+0.07·(1-0.10))=1.521;

[b]: TE=NDE·NIE.

Table 7: Simulation results: exact OR versus approximate OR estimators with OR as reference measure

		Exact OR results					**Approximate OR results**				
Effect	**True OR**[a]	Mean	Bias	SD	RMSE	CP	Mean	Bias	SD	RMSE	CP
Scenario 1: n=2000											
NIE	**1.056**	1.056	<0.001	0.026	0.026	0.951	1.062	0.005	0.029	0.030	0.953
NDE	**1.586**	1.600	0.015	0.261	0.261	0.938	1.620	0.034	0.264	0.266	0.938
TE[b]	**1.674**	1.690	0.015	0.275	0.276	0.940	1.719	0.045	0.283	0.286	0.941
Scenario 1: n=10,000											
NIE	**1.056**	1.056	<0.001	0.011	0.012	0.938	1.061	0.005	0.013	0.014	0.941
NDE	**1.586**	1.582	-0.003	0.105	0.105	0.962	1.602	0.017	0.106	0.107	0.964
TE[b]	**1.674**	1.671	-0.003	0.110	0.110	0.962	1.700	0.026	0.114	0.117	0.962
Scenario 2: n=2000											
NIE	**1.332**	1.346	0.015	0.173	0.173	0.940	1.418	0.087	0.198	0.216	0.928
NDE	**1.399**	1.411	0.013	0.229	0.229	0.949	2.650	1.252	0.865	1.521	0.438
TE[b]	**1.863**	1.893	0.030	0.348	0.350	0.947	3.763	1.901	1.339	2.325	0.492
Scenario 2: n=10,000											
NIE	**1.332**	1.334	0.002	0.075	0.075	0.940	1.407	0.076	0.089	0.117	0.874
NDE	**1.399**	1.396	-0.003	0.101	0.101	0.947	2.497	1.099	0.341	1.150	0.006
TE[b]	**1.863**	1.861	-0.001	0.163	0.163	0.927	3.519	1.656	0.554	1.747	0.010

Abbreviations: CP, coverage probability; NDE, natural direct effect; NIE, natural indirect effect; OR, odds ratio; RMSE, root mean square error; SD, standard deviation; TE, total effect.

[a]: True ORs are calculated from Table 1;

[b]: TE=NDE·NIE.

Appendix

Building and expanding on Online Appendix in Valeri and VanderWeele (2013), p.17-18, we show in the sequel how approximate expressions (3-4) can be obtained from exact expressions (8-12) in two steps as follows. In the first step, the rare-disease assumption (RDA) is used to replace (8) and (9) by

$$\frac{P(Y(1,M(0))=1|\boldsymbol{C}=\boldsymbol{c})}{P(Y(0,M(0))=1|\boldsymbol{C}=\boldsymbol{c})}$$

and

$$\frac{P(Y(1,M(1))=1|\boldsymbol{C}=\boldsymbol{c})}{P(Y(1,M(0))=1|\boldsymbol{C}=\boldsymbol{c})},$$

respectively. In the second step, the exact expressions (10-12) for $P(Y(1,M(1))=1|\boldsymbol{C}=\boldsymbol{c})$, $P(Y(1,M(0))=1|\boldsymbol{C}=\boldsymbol{c})$ and $P(Y(0,M(0))=1|\boldsymbol{C}=\boldsymbol{c})$ are replaced by

$$\underbrace{exp(\theta_0+\theta_1+\theta_2+\theta_3+\boldsymbol{\theta}_4'\boldsymbol{c})}_{\textbf{simplified by RDA}}\cdot\frac{exp(\beta_0+\beta_1+\boldsymbol{\beta}_2'\boldsymbol{c})}{1+exp(\beta_0+\beta_1+\boldsymbol{\beta}_2'\boldsymbol{c})}$$
$$+\underbrace{exp(\theta_0+\theta_1+\boldsymbol{\theta}_4'\boldsymbol{c})}_{\textbf{simplified by RDA}}\cdot\frac{1}{1+exp(\beta_0+\beta_1+\boldsymbol{\beta}_2'\boldsymbol{c})},$$

$$\underbrace{exp(\theta_0+\theta_1+\theta_2+\theta_3+\boldsymbol{\theta}_4'\boldsymbol{c})}_{\textbf{simplified by RDA}}\cdot\frac{exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})}{1+exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})}$$
$$+\underbrace{exp(\theta_0+\theta_1+\boldsymbol{\theta}_4'\boldsymbol{c})}_{\textbf{simplified by RDA}}\cdot\frac{1}{1+exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})},$$

and

$$\underbrace{exp(\theta_0+\theta_2+\boldsymbol{\theta}_4'\boldsymbol{c})}_{\textbf{simplified by RDA}}\cdot\frac{exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})}{1+exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})}+\underbrace{exp(\theta_0+\boldsymbol{\theta}_4'\boldsymbol{c})}_{\textbf{simplified by RDA}}\cdot\frac{1}{1+exp(\beta_0+\boldsymbol{\beta}_2'\boldsymbol{c})},$$

respectively.

Table A1: Inclusion of covariates in small and large set analyses

	Full set 1 (ICS exposure)	**Full set 2 (PA exposure)**	**Reduced set**
Number of included variables	27	27	4
Mother's and baby's characteristics			
Maternal age at the beginning of pregnancy (< 18, $>$ 18-34, $>$ 34 years)	X	X	X
Baby's sex (male/female)	X	X	X
Receipt of social assistance in the year before or during pregnancy (yes/no)	X	X	
Urban residency at delivery (yes/no)	X	X	
Maternal chronic conditions in the year before or during pregnancy			
Antiphospholipid syndrome (yes/no)	X	X	
Chronic hypertension (yes/no)	X	X	
Diabetes mellitus (yes/no)	X	X	X
Cystic fibrosis of the pancreas (yes/no)	X	X	
Uterine defects (yes/no)	X	X	
Pregnancy-related variables			
Gestational diabetes (yes/no)	X	X	X
Eclampsia/pre-eclampsia (yes/no)	X	X	
Anaemia during pregnancy (yes/no)	X	X	
Placental conditions[a] (yes/no)	X	X	
Placental abruption (yes/no)	X	n/a	
Vaginal bleeding (yes/no)	X	X	
Maternal infections during pregnancy [b] (yes/no)	X	X	
Fetal-maternal hemorrhage (yes/no)	X	X	
Pregnancy-induced hypertension (yes/no)	X	X	
Use of beta-blockers during pregnancy (yes/no)	X	X	
Asthma-related variables			
ICS during pregnancy (yes/no)	n/a	X	
Leukotriene-receptor antagonists during pregnancy (yes/no)	X	X	
SABA during pregnancy (0, $>$ 0-3, $>$3 doses/week)	X	X	
Oral corticosteroids during pregnancy (yes/no)	X	X	
Intranasal corticosteroids during pregnancy (yes/no)	X	X	
$\geq$ 1 ED visit for asthma during pregnancy (yes/no)	X	X	
LABA during pregnancy (yes/no)	X	X	
$\geq$ hospitalization for asthma during pregnancy (yes/no)	X	X	
Severity of asthma in the year before conception (mild, moderate, severe)	X	X	

Abbreviations: ED, emergency department; ICS, inhaled corticosteroids; LABA, long-acting beta2-agonists; PA, placental abruption; SABA, short-acting beta2-agonists.

[a]: Single umbilical artery, velamentous umbilical cord insertion, bilobate placenta, suboptimal implantation site, placenta previa, placental anomalies, and cord anomalies: yes, if at least one condition is present; no, otherwise;

[b]: Urinary-genital infections, malaria, trypanosomiasis, cytomegalovirus, toxoplasmosis, rubella, herpes virus: yes, if at least one condition is present; no, otherwise;

References

Ananth, C. V. and VanderWeele, T. J. (2011). Placental abruption and perinatal mortality with preterm delivery as a mediator: disentangling direct and indirect effects. *American Journal of Epidemiology*, 174(1):99–108.

Casella, G. and Berger, R. (2002). *Statistical Inference*. Duxbury Thomson Learning, 2nd edition.

Cassell, D. L. (2007). Don't be loopy: Re-sampling and simulation the SAS® Way. In *Proceedings of the SAS Global Forum*.

Chernick, M. R. (2011). *Bootstrap methods: A guide for practitioners and researchers*, volume 619. John Wiley & Sons, 2nd edition.

Cummings, P. (2009). The relative merits of risk ratios and odds ratios. *Archives of Pediatrics and Adolescent Medicine*, 163(5):438–445.

Daniel, R. M., De Stavola, B. L., and Cousens, S. N. (2011). gformula: Estimating causal effects in the presence of time-varying confounding or mediation using the g-computation formula. *Stata Journal*, 11(4):479–517.

Fang, J. (2011). Using SAS procedures FREQ, GENMOD, LOGISTIC, and PHREG to estimate adjusted relative risks –A case study. In *Proceedings of the SAS Global Forum*.

Imai, K., Keele, L., and Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, 25(1):51–71.

Localio, A. R., Margolis, D. J., and Berlin, J. A. (2007). Relative risks and confidence intervals were easily computed indirectly from multivariable logistic regression. *Journal of Clinical Epidemiologyl*, 60(9):874–882.

Pearl, J. (2001). Direct and indirect effects. In *Proceedings of the seventeenth conference on uncertainty in artificial intelligence*, pages 411–420. Morgan Kaufmann Publishers Inc.

Robins, J. M. and Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, 3(2):143–155.

Samoilenko, M., Blais, L., Cossette, B., Forget, A., and Lefebvre, G. (2016). Assessing the dose-response relationship between maternal use of inhaled corticosteroids therapy and birth weight: a generalized propensity score approach. *Observational Studies*, (Article 8):90–118.

Starkopf, L., Andersen, M. P., Gerds, T. A., Torp-Pedersen, C., and Lange, T. (2017). Comparison of five software solutions to mediation analysis. Research report, University of Copenhagen.

Steen, J., Loeys, T., Moerkerke, B., and Vansteelandt, S. (2017). Medflex: an R package for flexible mediation analysis using natural effect models. *Journal of Statistical Software*, 76(11).

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., and Imai, K. (2014). Mediation: R package for causal mediation analysis. *Journal of Statistical Software*, 59(5).

Valeri, L. and VanderWeele, T. J. (2013). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*, 18(2):137–150.

VanderWeele, T. J. (2015). *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press.

VanderWeele, T. J. (2016). Mediation analysis: A practitioner's guide. *Annual Review of Public Health*, 37:17–32.

VanderWeele, T. J. and Vansteelandt, S. (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and Its Interface*, 2:457–468.

VanderWeele, T. J. and Vansteelandt, S. (2010). Odds ratios for mediation analysis for a dichotomous outcome. *American Journal of Epidemiology*, 172(12):1339–1348.

VanderWeele, T. J. and Vansteelandt, S. (2014). Mediation analysis with multiple mediators. *Epidemiologic Methods*, 2(1):95–115.

Vansteelandt, S., Bekaert, M., and Lange, T. (2012). Imputation strategies for the estimation of natural direct and indirect effects. *Epidemiologic Methods*, 1(1):131–158.

Wang, A. and Arah, O. A. (2015). G-computation demonstration in causal mediation analysis. *European Journal of Epidemiology*, 30(10):1119–1127.