Conditioning on Intermediates in Perinatal Epidemiology

Tyler J. VanderWeele,[a] *Sunni L. Mumford,*[b] *and Enrique F. Schisterman*[b]

Abstract: It is common practice in perinatal epidemiology to calculate gestational-age-specific or birth-weight-specific associations between an exposure and a perinatal outcome. Gestational age or birth weight, for example, might lie on a pathway from the exposure to the outcome. This practice of conditioning on a potential intermediate has come under critique for various reasons. First, if one is interested in assessing the overall effect of an exposure on an outcome, it is not necessary to stratify, and indeed, it is important not to stratify, on an intermediate. Second, if one does condition on an intermediate, to try to obtain what might conceived of as a "direct effect" of the exposure on the outcome, then various biases and paradoxical results can arise. It is now well documented theoretically and empirically that, when there is an unmeasured common cause of the intermediate and the outcome, associations adjusted for the intermediate are subject to bias. In this paper, we propose 3 approaches to facilitate valid inference when effects conditional on an intermediate are in view. These 3 approaches correspond to (i) conditioning on the predicted risk of the intermediate, (ii) conditioning on the intermediate itself in conjunction with sensitivity analysis, and (iii) conditioning on the subgroup of individuals for whom the intermediate would occur irrespective of the exposure received. The second and third approaches both require sensitivity analysis, and they result in a range of estimates. Each of the 3 approaches can be used to resolve the "birth-weight paradox" that exposures such as maternal smoking seem to have a protective effect among low-birth-weight infants. The various methodologic approaches described in this paper are applicable to a number of similar settings in perinatal epidemiology.

(*Epidemiology* 2012;23: 1–9)

Submitted 29 September 2011; accepted 30 September 2011.

From the [a]Departments of Epidemiology and Biostatistics, Harvard School of Public Health, Boston, MA; and [b]Epidemiology Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD.

Supported by National Institutes of Health grant HD060696 (to T.J.V.) and Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health (to S.L.M. and E.F.S.). The authors reported no other financial interests related to this research.

Editors' note: *Related articles appear on pages 10 and 13.*

Correspondence: Tyler J. VanderWeele, Departments of Epidemiology and Biostatistics, Harvard School of Public Health, 677 Huntington Ave, Boston, MA 02115. E-mail: tvanderw@hsph.harvard.edu.

ISSN: 1044-3983/12/2301-0001
DOI: 10.1097/EDE.0b013e31823aca5d

Conditioning on an intermediate is of concern in all areas of epidemiologic research.[1–3] In particular, there is tension in the perinatal epidemiology literature between the desire to obtain gestational-age-specific or birth-weight-specific associations[4–6] and the increasing awareness that conditioning on such variables can give rise to severe biases.[7–10] The difficulty arises because gestational age and birth weight may be on a pathway from the exposure of interest to the perinatal outcome or may be descendents of such a variable. For example, if the exposure were smoking and the outcome were infant mortality, smoking may in part affect infant mortality through its effects on fetal growth or the timing of delivery. When one conditions on such an intermediate, without also controlling for the common causes of the intermediate and the outcome, biased results and paradoxical findings can emerge.[1–3,11]

It has, for example, long been documented[12–14] that among infants with the lowest birth weight, maternal smoking seems to have a protective effect on infant mortality. This seemingly perplexing association is often referred to as the "birth-weight paradox." The apparent protective effect of maternal smoking is an artifact of conditioning on an intermediate without adequate control for intermediate-outcome confounding.[7] In the case of the birth-weight paradox, birth defects may be a common cause of both birth weight and infant mortality (Fig.). Birth defects are not controlled for in the analysis. For mothers who do smoke and have low-birth-weight infants, low birth weight might be a consequence of either smoking or of a birth defect. For mothers who do not smoke and have low-birth-weight infants, low birth weight cannot be a consequence of smoking and so some other cause must be operating.[7] A comparison of smoking and nonsmoking mothers, without controlling for birth defects, will artificially bias the comparison because for this group of low-birth-weight infants, not smoking and low birth weight together is more likely indicative of the presence of a birth defect. This form of bias is sometimes referred to as collider-stratification bias and is described in more detail elsewhere.[15–18] Similar issues arise, sometimes in less severe form, throughout perinatal epidemiology.[19] The intermediate variable need not have an effect on the outcome for such biases to arise; all that is necessary is for the exposure to affect the intermediate and for there to be an unmeasured common cause of the intermediate and the outcome.

A considerable literature has developed pointing out this problem of conditioning on an intermediate. However, little of that literature[8,9,20] has offered solutions other than

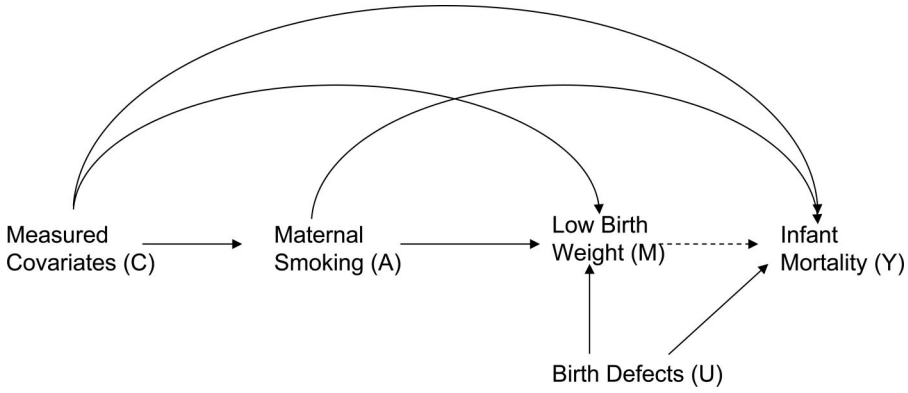

FIGURE. Diagram illustrating relationships between an exposure, eg, smoking (*A*); an intermediate, eg, low birth weight (*M*); an outcome, eg, infant mortality (*Y*); and both measured (*C*) and unmeasured (*U*) confounders.

abandoning conditioning on the intermediate altogether. Simply not conditioning on the intermediate will often be the correct way to proceed with analysis. If the overall effect of the exposure on the outcome is of interest, then there is no reason to condition on an intermediate. Conditioning on an intermediate in general will be of interest only when other types of effect, such as the direct effect of the exposure on the outcome (not through the intermediate) are in view. We will return at the end of the paper to these questions of whether and when to condition on an intermediate. However, it is important to emphasize that this choice should be determined by first clarifying the effect of interest.

In this paper, we discuss 3 analytic approaches to help draw inferences when the effect of interest is such that it may be obtained by conditioning on an intermediate. The 3 approaches include conditioning on the predicted risk of the intermediate, conditioning on the intermediate itself with sensitivity analysis, and conditioning on the principal stratum. Each approach carries with it a different interpretation, a different set of assumptions, and, when relevant, different methods for sensitivity analysis. When sensitivity analysis techniques are used, often a range of estimates, rather than a single estimate, will be obtained. As will be observed later, the 3 approaches estimate different causal effects, and the resulting estimates would not be expected to all be the same. We illustrate each by application to data exemplifying the birth-weight paradox; with the application of each of the 3 approaches we describe, the birth-weight paradox will be seen to dissolve. The approaches are applicable to a variety of similar settings in perinatal epidemiology.

Notation and Definitions

We use notions of counterfactuals (or "potential outcomes") through parts of this paper.[21,22] We will let A denote our exposure of interest (eg, smoking), M will be the intermediate (eg, low birth weight, defined as less than 2500 g), and Y will be outcome (eg, infant mortality). We will let C denote some set of baseline characteristics measured prior to or concurrent with the exposure. The relationships among the variables are depicted in the Figure. The approaches we describe later will be applicable irrespective of whether the intermediate has an actual effect on the outcome.

We will let Y_a denote the counterfactual outcome for each individual if the exposure had been set to level a and let M_a denote intermediate if the exposure had been set to level a. Thus, if A were binary, for each individual we would have 2 possible counterfactual outcomes Y_1 and Y_0, corresponding to what would have happened to the individual (eg, with or without smoking); and likewise we will have 2 possible counterfactual intermediates M_1 and M_0. For each individual, we will be able to observe only one of Y_1 or Y_0, corresponding to the exposure that was in fact received; and likewise for M_1 and M_0. We will say that the effect of A on Y is unconfounded conditional on C if the groups with $A = 1$ and $A = 0$ are comparable in their distribution of potential outcomes (Y_0, Y_1) conditional on C. Likewise, we will say that the effect of A on M is unconfounded conditional on C if the groups with $A = 1$ and $A = 0$ are comparable in their distribution of potential outcomes (M_0, M_1) conditional on C.

NCHS Data Illustrating the Birth Weight Paradox

We use cohort-linked birth certificate and infant mortality files for year 1997 from the National Center for Health Statistics (NCHS). These are complete files for all US births in year 1997 with 1-year follow-up for infant mortality. Only singletons were included in the analysis. Demographic characteristics for this sample are given in Table 1. Smoking status was dichotomized (any smoking during pregnancy vs. none). Low birth weight was defined as birth weight less than 2500 g. Covariates include sex (boy/girl), maternal race (black, white, other), maternal education (less than high school, high school, more than high school), maternal age (<15, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–54), nulliparous (yes/no), marital status (married/single), and prior preterm birth (yes/no). Table 2 gives infant mortality statistics stratified by smoking and low-birth-weight status. The crude odds ratio (OR) relating infant mortality and smoking is 1.69 (95% confidence interval [CI]: 1.63 to 1.75). The odds ratio relating infant mortality and smoking adjusted for the covariates C using logistic regression (omitting birth weight)

TABLE 1. Baseline Characteristics of Women With Singleton Pregnancies in the 1997 Cohort-linked Birth Certificate Infant Mortality Files From the National Center for Health Statistics, by Birth Weight and Smoking Status

		Birth Weight[a]			
	Overall (n = 3,773,369) No. (%)	**<2500 g (n = 229,863) No. (%)**	**≥2500 g (n = 3,541,919) No. (%)**	**Smoking Status[b] (n = 397,780) No. (%)**	**Nonsmoker (n = 2,606,836) No. (%)**
Infant mortality					
Live birth	3,749,676 (99)	215,667 (94)	3,532,669 (100)	393,830 (99)	2,591,452 (99)
Infant death	23,693 (1)	14,196 (6)	9250 (0.3)	3950 (1)	15,384 (1)
Maternal age (years)					
<15	10,011 (0.3)	1285 (1)	8724 (0.2)	665 (0.2)	7607 (0.3)
15–19	476,087 (13)	41,113 (18)	434,766 (12)	68,699 (17)	320,588 (12)
20–24	923,143 (25)	58,558 (26)	864,201 (24)	123,929 (31)	620,988 (24)
25–29	1,040,805 (28)	54,121 (24)	986,280 (28)	97,861 (25)	733,045 (28)
30–34	855,150 (23)	44,099 (19)	810,734 (23)	67,067 (17)	604,523 (23)
35–39	392,308 (10)	24,634 (11)	367,456 (10)	33,695 (9)	268,936 (10)
40–44	72,960 (2)	5784 (3)	67,124 (2)	5698 (1)	49,213 (2)
45–54	2824 (0.1)	261 (0.1)	2561 (0.1)	163 (0.04)	1881 (0.1)
Missing	0 (0)	8 (0.003)	73 (0.002)	3 (0.001)	55 (0.002)
Maternal race					
White	2,987,619 (79.2)	151,198 (66)	2,835,430 (80)	337,702 (85)	2,014,565 (77)
Black	581,504 (15.4)	66,241 (29)	514,951 (15)	50,280 (13)	467,211 (18)
Other	204,246 (5.4)	12424 (5)	191,538 (5)	9798 (3)	125,060 (5)
Maternal education					
< High school	829,668 (22)	65,811 (29)	763,588 (22)	140,321 (35)	483,145 (19)
High school	1,225,534 (33)	80,130 (35)	1,144,962 (32)	170,306 (43)	824,000 (32)
> High school	1,663,457 (45)	78,703 (34)	1,584,384 (45)	82,006 (21)	1,264,335 (49)
Missing	0 (0)	5219 (2)	48,985 (1)	5147 (1)	35,356 (1)
Married	1,229,086 (33)	108,636 (47)	1,119,828 (32)	209,813 (53)	772,515 (30)
Nulliparous	1,261,874 (33)	87,202 (38)	1,174,241 (33)	102,694 (26)	902,769 (35)
Missing	28,167 (1)	2331 (1)	25,237 (1)	2641 (1)	18,958 (1)
Prior preterm birth	44,177 (1)	9684 (4)	34,474 (1)	8861 (2)	25,506 (1)
Missing		3355 (2)	41,639 (1)	3854 (1)	21,501 (1)
Infant male sex	1,931,390 (51)	109,263 (48)	1,821,269 (51.4)	203,461 (51.2)	1,335,443 (51.2)

[a]Missing information on birth weight for 1587 women.
[b]Missing information on smoking status for 768,753 women.

TABLE 2. Infant Mortality (No. Deaths per 1000 Live Births) Among Women With Singleton Pregnancies in the 1997 Cohort-linked Birth Certificate Infant Mortality Files From the National Center for Health Statistics, by Birth Weight and Smoking Status

	Birth Weight ≥2500 g[a]		**Birth Weight <2500 g[b]**		**Overall[c]**		
	Smoker	**Nonsmoker**	**Smoker**	**Nonsmoker**	**Smoker**	**Nonsmoker**	**Total**
Infant mortality							
Live birth	353,335	2,453,633	40,383	137,154	393,830	2,591,452	3,749,676
Infant death	1729	5838	2192	9387	3950	15,384	23,693
Deaths per 1000	4.9	2.4	51.5	64.1	2.4	5.9	6.3

[a]Missing information on 40,747 women.
[b]Missing information on 727,384 women.
[c]Missing information on 768,753 women.

is 1.53 (1.47 to 1.59), which would be the more appropriate analysis for assessing the overall effect of smoking on infant mortality.

Finally, we use a logistic regression model of infant mortality Y on smoking A, low birth weight M, an $A \times M$ interaction, and covariates C:

$$\text{log it}\,(P(Y = 1|a, m, c)) = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4' c. \quad (1)$$

We can then calculate the odds ratio comparing smoking versus nonsmoking for both low-birth-weight infants (OR = e^{θ_1}) and normal-birth-weight infants (OR = $e^{\theta_1+\theta_3}$). Using the NCHS data for infants weighing more than 2500 g, we obtain an adjusted odds ratio of 1.87 (1.76 to 1.97). For low-birth-weight infants, we obtain an adjusted odds ratio of 0.76 (0.72 to 0.80). One might interpret these finding as evidence for an interaction between birth weight and smoking.[6] In any case, we see that smoking seems to have a protective effect for low-birth-weight infants.

Approach 1. Conditioning on the Risk of an Intermediate

As noted earlier, when we calculate effect measures conditional on intermediate variables, biases can result. However, such problems do not in general arise in a similar manner when conditioning on baseline covariates. More specifically, if the effect of A on Y is unconfounded conditional on baseline covariates C, then we can validly estimate the causal effect of A on Y conditional on $C = c$ simply by calculating our effect measures within strata of C.

The variables C themselves might be predictive of the intermediate M. Our first approach to "conditioning on an intermediate" will consist of conditioning on the "risk of the intermediate" predicted by the baseline covariates, rather than conditioning on the intermediate itself. Although this will not allow us to get birth-weight-specific associations, it will at least allow us to assess whether smoking is protective among the group of infants who are most at-risk for being of low birth weight.

More specifically let $P(M = 1|C = c)$ denote the probability of low birth weight conditional on covariates $C = c$. Predicted probabilities for each individual could be obtained by a logistic regression of low birth weight M on the covariates C:

$$\text{log it}\,(P(M = 1|c)) = \delta_0 + \delta_1' c.$$

Once we have estimated the parameters (δ_0, δ_1') from the logistic regression, we can calculate predicted probabilities of low birth weight based on each woman's covariate values c. Note that this predicted probability is a function only of the woman's baseline covariates c, and we can thus condition on it in the analysis of infant mortality because we are only conditioning on a function of the baseline covariates; we are not in fact conditioning on the intermediate. We now define a new high-risk-of-low-birth-weight variable H. We will let $H = 1$ for mothers who have predicted probabilities of having low-birth-weight infants above the 95th percentile, and we let $H = 0$ otherwise. Note that our high-risk variable H is again simply a function of the covariate values c. To examine the association between smoking and infant mortality among infants who are or are not high risk for low birth weight, we could fit a logistic regression model of infant mortality Y on smoking A, high-risk-status H, an $A \times H$ interaction, and covariates C:

$$\text{log it}\,(P(Y = 1|a, h, c)) = \lambda_0 + \lambda_1 a + \lambda_2 h + \lambda_3 ah + \lambda_4' c. \quad (2)$$

We can then calculate the odds ratio comparing smoking versus nonsmoking for low-risk infants (OR = e^{λ_1}) and high-risk infants (OR = $e^{\lambda_1+\lambda_3}$). Because our risk-of-low-birth-weight variable H is essentially a function just of the baseline covariates C, we no longer run into the same problems of conditioning on an intermediate (eg, collider-stratification bias). However, the interpretation of the estimate is now different. The estimate captures the overall effect of smoking on infant mortality (including that through the intermediate) for those at high risk of the intermediate. This estimate from approach 1 thus does not capture the effect of the exposure that is not through the intermediate (the "direct effect"); that will be the focus of approach 2 below. Approach 1 does not give conclusions about direct effects. To determine birth weight risk categories, we apply this approach to the NCHS data using the sex, maternal race, maternal education, maternal age, nulliparity, marital status, and prior preterm birth. For low-risk infants ($H = 0$), we obtain an adjusted odds ratio of 1.55 (95% CI = 1.49 to 1.62). For high-risk infants, we obtain an adjusted odds ratio of 1.30 (1.16 to 1.46). These are the effects of smoking for those with lower versus higher predicted probabilities of low birth weight, respectively. The odds ratio for smoking is lower for high-risk infants than for low-risk infants (1.30 vs. 1.55), which is perhaps not particularly surprising as high-risk infants will generally have higher baseline risk without smoking. However, even for high-risk infants, it is not the case that smoking has an apparent protective effect on infant mortality; it has a harmful effect with an odds ratio for mortality of 1.30. By not conditioning on the intermediate (but rather on a set of baseline covariates that predict low birth weight), we have avoided the problems associated with conditioning on an intermediate and have circumvented the birth-weight paradox. Similar analyses could be done with other percentile cutoffs to define "high-risk." A limitation of this approach is that a larger set of baseline covariates with better predictive power of the intermediate could change the estimates obtained; the "low-risk" and "high-risk" designations are relative to the covariates being used. In some cases, high-risk may not be a reasonable label if, for example, even those above the 95th percentile still have relatively low predicted probability of the intermediate.

Approach 2. Conditioning on an Intermediate With Sensitivity Analysis

Consider again the logistic regression in equation (1) in which we condition on the intermediate itself, *M,* (eg, gestational age or low birth weight). Our odds ratios will be biased because of unmeasured confounding of the intermediate-outcome relationship (eg, by birth defects or malnutrition). One approach to address these biases is the use of sensitivity analysis to assess how such an unmeasured common cause U of the intermediate and the outcome might affect our odds ratio estimates. Specifically, suppose U is a binary variable indicating the presence of a common cause of the intermediate, say low birth weight and the outcome, infant mortality. Let γ denote the odds of infant mortality Y comparing $U = 1$ and $U = 0$ conditional on smoking exposure A, low-birth-weight status M, and the covariates C, and let π_{am} denote the prevalence of U among those with smoking status a ($a = 1$ or $a = 0$) and low-birth-weight status m ($m = 1$ or $m = 0$). We have shown elsewhere[3] that if the outcome is rare, so that odds ratios approximate risk ratios, and if U increases the risk of infant mortality Y by the same factor for low-birth-weight and normal-birth-weight infants, then the ratio between the estimate not controlling for U and the estimate that would have been obtained after controlling for U is given by the following:

$$B = \frac{1 + (\gamma - 1)\pi_{1m}}{1 + (\gamma - 1)\pi_{0m}}. \quad (3)$$

The corrected odds ratio (that would have been obtained when adjusting for U) can then be calculated by dividing the estimated odds ratio by the bias factor B. Note that this ratio may differ for low-birth-weight infants ($m = 1$) and normal-birth-weight infants ($m = 0$) because the prevalences of U, π_{1m}, and π_{0m}, may differ for normal and low-birth-weight infants. One could alternatively explore potential unmeasured confounding through simulations.[9,20]

Suppose we were to apply this approach to the NCHS data. For low-birth-weight infants, we had obtained a (potentially biased) odds ratio of 0.76 (95% CI = 0.72 to 0.80). Suppose now that for low-birth-weight infants the prevalence of U for smoking mothers is $\pi_{1m} = 0.025$ but that the prevalence of U for nonsmoking mothers is $\pi_{0m} = 0.14$ (because if smoking were not the cause of low birth weight, this renders some other explanation/cause more likely). If the effect of U on infant mortality were a 3.5-fold increase ($\gamma = 3.5$), we would have a bias factor in equation (3) of 0.79 (1.06/1.35 = 0.72) and thus a corrected odds ratio of 0.76/0.79 = 0.96 (0.91 to 1.01). If the effect of U were instead a 5-fold increase ($\gamma = 5$), we would have a bias factor of 0.59 (1.3/2.2 = 0.59) and a corrected odds ratio of 0.76/0.59 = 1.29 (95% CI: 1.22 to 1.36).

We have selected our sensitivity in part based on estimates for birth defects obtained from the literature or transformations of estimates in the literature.[23–25] Based on estimates for the fraction of infant mortality due to birth defects, the sensitivity analysis parameter γ may be even higher.[23] However, in general, sensitivity analysis techniques can be helpful even in the absence of such knowledge. One can assess how large the effects would have to be to explain away the result, and readers can then decide whether they think such values are or are not plausible. One can also specify several different sets of sensitivity analysis parameters and obtain a range of different estimates. No specific set of parameters needs to be believed for such a sensitivity analysis to be informative. In some cases, only a little confounding may be necessary to explain away the effect; in others, quite a lot. These techniques can also in principle be used with multiple unmeasured confounding variables.[3] Here, we have considered birth defects, but malnutrition may also be a common cause. Analyses that considered jointly the consequences of birth defects and malnutrition would even more easily be able to explain away the observed association. It should also be noted that the sensitivity analysis used here assumed that the unmeasured confounder U increased the risk of infant mortality Y by the same factor for low-birth-weight and normal-birth-weight infants; this assumption can also be relaxed.[3] However, if a sensitivity analysis under this simplifying assumption explains away the association, then sensitivity analysis dropping this assumption would still explain away the effect estimate. This is because the sensitivity parameter values that explained away the estimate under the simplified technique would still be within the space of sensitivity parameter values under the more general technique.

Similar analyses could be used to adjust the odds ratio for normal-birth-weight infants. However, we see here that by using the bias formula to attempt to correct for possible unmeasured confounding of the birth-weight-mortality relationship, our corrected odds ratios for smoking among low-birth-weight infants rise above 1 for plausible values of the sensitivity analysis parameters. The birth-weight paradox vanishes: maternal smoking no longer seems to have a protective effect among low-birth-weight infants. Thus, although estimates conditioning on an intermediate may be biased due to unmeasured confounding of the intermediate-outcome relationship, sensitivity analysis in conjunction with these biased estimates can allow for reasoning about the estimates that would have been obtained had it been possible to control for such unmeasured confounders.

We note that in this and similar examples, if it were the case that the measured covariates C and the unmeasured confounder U together sufficed to control for confounding of the joint effects of smoking and low birth weight on infant mortality, then the corrected odds ratios could be interpreted as controlled direct effects.[1,3,26] In the causal inference liter-

ature, controlled direct effects are usually conceived of as the effect of the exposure under interventions to fix the intermediate.[1,26] In some settings, direct interventions to fix the intermediate (eg, gestational age or birth weight) to a particular value may be implausible, and the analysis is perhaps best conceived of as one of adjusted association. To avoid confounded associations conditional on the intermediate, one controls for measured confounding by analytic adjustment and unmeasured confounding by sensitivity analysis. VanderWeele and Hernández-Diaz[27] apply a similar approach to that described here in the context of preeclampsia as the exposure, preterm birth as the intermediate, and cerebral palsy as the outcome, to examine whether there is a direct effect of preeclampsia on cerebral palsy not through preterm birth.

Approach 3. Conditioning on Principal Stratum

Thus far we have considered conditioning on the risk of the intermediate being present, and on the intermediate itself. Our third approach involves assessing the effect of the exposure on the outcome among the subpopulation for whom the intermediate would be present irrespective of exposure status. For example, we might be interested in the effect among the subpopulation that would be low birth weight irrespective of maternal smoking. This subgroup for whom the intermediate will occur irrespective of the exposure is sometimes referred to as a "principal stratum."[28] More generally, a principal stratum is a subgroup defined by the joint potential outcomes (M_0, M_1). If the exposure A and the intermediate M are binary, then there are 4 principal strata: those for whom the intermediate will not occur irrespective of exposure status ($M_0 = 0, M_1 = 0$, "never low birth weight"), those for whom the intermediate will occur with exposure but not without ($M_0 = 0, M_1 = 1$, "low birth weight only with smoking"), those for whom the intermediate will occur without the exposure but not with ($M_0 = 1, M_1 = 0$, "low birth weight only if nonsmoking"/"defiers"), and those for whom the intermediate will occur irrespective of exposure status ($M_0 = 1, M_1 = 1$, "always low birth weight"). If we are interested in whether smoking has a protective effect among low-birth-weight infants, one potentially relevant question to ask within this context of principal stratification is whether smoking has a protective effect among the subpopulation who would be low birth weight irrespective of exposure status ($M_0 = 1, M_1 = 1$). In counterfactual notation this effect is given by the following:

$$PSDE = E[Y_1 - Y_0 | M_0 = 1, M_1 = 1].$$

This is sometimes referred to as a "principal-stratum direct effect."[29,30]

The advantage of this approach using principal stratification is that, like the first approach, we essentially avoid the problem of conditioning on the intermediate directly. Instead we condition on the principal stratum, which is essentially an underlying characteristic of the individual. It is like conditioning on a baseline covariate.

The disadvantage of this approach is that we do not know who is in each principal stratum. For example, we do not know which infants will be low birth weight irrespective of maternal smoking. Because we cannot identify the individuals who fall in each principal stratum, we cannot estimate the principal strata direct effect above directly from the data. However, one can attempt to assess the magnitude of this effect by using sensitivity analysis techniques for principal strata.[3,30–32] Specifically, it has been shown[32] that if one calculates the crude outcome difference between the exposed ($A = 1$) and the unexposed ($A = 0$) among those for whom the intermediate is in fact present ($M = 1$):

$$E[Y|A = 1, M = 1] - E[Y|A = 0, M = 1]$$

then the principal stratum direct effect can be expressed as the difference between the crude outcome difference and a sensitivity analysis parameter, under an assumption that there are no individuals for whom the intermediate would occur if unexposed, but not if exposed (ie, no defiers). In the context of the smoking-birth weight example, this would imply that there are no infants who would be normal birth weight if their mother smoked but who would be low birth weight if their mother did not smoke. If this is the case, then it can be shown that[32]:

$$PSDE = E[Y|A = 1, M = 1] - E[Y|A = 0, M = 1] - \alpha,$$

where the sensitivity analysis parameter α is given by the following:

$$\alpha = E[Y_1|A = 1, M = 1] - E[Y_1|A = 0, M = 1].$$

The interpretation of this sensitivity analysis parameter α is the difference in infant mortality rates under maternal smoking for 2 populations, the population for whom the infant would be low birth weight if the mother smoked and the population for whom the infant would be low birth weight if the mother did not smoke. Because this second population consists of those who would be low birth weight even if the mother did not smoke, it is probably a less healthy population that is likely to have a higher infant mortality rate if the mother smokes. The parameter α will thus likely be negative. The parameter is not identified from the data; rather, an investigator can specify different plausible values of this parameter to assess the principal-stratum direct effect. One will again obtain a range of estimates corresponding to the different sensitivity analysis specifications. Further detail on inference for principal-stratum direct effects can be found elsewhere.[3,30–34]

We now apply this principal-stratification approach to the NCHS data. If we first calculate the crude infant mortality

difference between the smoking and nonsmoking mothers among infants of low birth weight, we obtain $0.051 - 0.064 = -0.013$ (95% CI = -0.015 to -0.011) and once again with the crude analysis, it seems that smoking has a protective effect on infant mortality for low-birth-weight infants. If we want to calculate the principal-stratum direct effect (ie, the effect of smoking on infant mortality among the subpopulation who would have been low birth weight irrespective of smoking exposure), then we need to adjust this crude estimate by the sensitivity parameter α. If we thought that the difference in infant mortality rates under smoking when comparing the population for whom the infant would have been low birth weight if the mother smoked and the population for whom the infant would have been low birth weight even if the mother did not smoke were -0.02 we would then obtain an estimate of the principal-stratum direct effect of 0.007 (0.005 to 0.009); if we thought that this difference were -0.03 we would obtain an estimate of the principal-stratum direct effect of 0.017 (0.015 to 0.019). These estimates of the principal stratum direct effect would suggest that smoking has a harmful effect on the subpopulation who would be low birth weight irrespective of smoking exposure. The principal-stratum direct effect is positive; the birth-weight paradox is again resolved. A disadvantage of this approach is that, even if we use sensitivity analysis to assess the principal stratum direct effect, we still do not know which individuals are in this principal stratum (ie, we do not know which infants would be low birth weight irrespective of maternal smoking).

DISCUSSION

We have described 3 approaches related to calculating effects of an exposure, conditional on some intermediate variable or a variant of it. Each of the approaches has a unique interpretation, requires different assumptions, and has its own strengths and weaknesses. In the first approach, we conditioned on the risk of the intermediate as predicted by baseline covariates, rather than on the intermediate itself. By conditioning on the risk predicted by baseline covariates, one could still evaluate the effect of an exposure (eg, smoking) for high-risk subgroups, while avoiding the biases ordinarily associated with conditioning on an intermediate. The disadvantage of this first approach is that the effect one obtains is not specific to individuals for whom the intermediate (eg, low birth weight or gestational age) actually occurs, but for those at high risk of occurrence. If the intermediate is relatively rare, this may not be an accurate reflection of the effect of the exposure for those for whom the intermediate will in fact develop. Moreover, the effect that is captured is the overall effect of the exposure (including that through the intermediate) for those who are at high risk of the intermediate; this may not be what is of substantive interest.

In the second approach, we conditioned on the intermediate itself. The disadvantage of this approach is that it will induce bias whenever there is an unmeasured common cause of the intermediate and the outcome; one must then attempt to correct for the bias through sensitivity analysis. The advantage of this approach is that, after correction through sensitivity analysis, one obtains a range of estimates (according to the range of sensitivity analysis parameters specified) of the effect of the exposure on the outcome for individuals with the intermediate, corresponding to the direct effect of the exposure on the outcome not through the intermediate. Obtaining this direct effect is arguably what is often the goal when researchers condition on an intermediate in perinatal epidemiology.

In the third approach, one conditions on the subpopulation for whom the intermediate would occur irrespective of exposure. This approach has the advantage that this is a particularly high-risk group and it is a group for whom the intermediate will necessarily occur. The disadvantage of the third approach, like the second, is that naive estimators of this effect are biased and correction needs to be made through sensitivity analysis; one again will obtain a range of estimates. A further disadvantage with this third approach is that, even after applying the methodology, we still do not know who is in the subpopulation such that the intermediate will occur irrespective of the exposure. We believe that the second approach presented here will in general be the one of greatest interest. In essence, it captures the direct effect of the exposure on the outcome not through the intermediate, which we believe is the effect that is often desired when investigators condition on an intermediate.

In many studies, the overall effect of the exposure on the outcome may be of central interest and none of the approaches described here is then needed. The approaches here are of relevance only when the investigator is interested in the direct effect of the exposure not through the intermediate or in the effect of the exposure for certain groups at high risk for, or certain of having, the intermediate. In these cases, it may be desirable to employ all 3 approaches to develop a fuller understanding of the relationships among the exposure, intermediate, and the outcome. In some settings, the 3 effects may all be in a consistent direction. However, it is important to note that the 3 approaches need not all give effect estimates in the same direction. Having effect estimates in different directions for the 3 approaches is not necessarily an indication that one or more of the estimates is in the wrong direction. The 3 approaches estimate 3 different effects (effects for 3 different populations) and these may in fact be in different directions. Moreover, the effects under approaches 1 and 3 are overall effects (including the effect through the intermediate) for their respective subpopulations, whereas approach 2 is a direct effect, the effect not through the intermediate. In our analyses performed earlier, all 3 approaches, after sensitivity analysis, suggested a harmful effect of maternal smoking on infant mortality for (i) those at high

risk of low birth weight, (ii) those who in fact were low birth weight, and (iii) those who would be low birth weight irrespective of smoking exposure.

In some applications with an intermediate, one desires not simply to estimate the effects of the exposure conditional on the intermediate, but also to partition the total effect into the proportion through the intermediate and the proportion through other pathways. Such analyses are subject to an even wider array of possible confounding biases[26,35–37] and are beyond the scope of the present paper. Nevertheless, researchers have begun to develop methodology to allow for these analyses using so-called natural direct and indirect effects; sensitivity analysis can be used to address some of the biases that arise in these settings.[3,38,39] For an example of such an analysis in perinatal epidemiology, see the study of Ananth and VanderWeele[40] that examines the extent to which the effect of placental abruption on perinatal mortality outcomes is mediated through premature birth.

When one does condition on an intermediate (or variants thereof under the approaches described earlier), it is important to be clear about what the scientific or policy question is. The approaches we have described may be of service when one is interested in assessing the presence of a direct effect, or when one is interested in the effect of an exposure among particularly vulnerable subpopulations. However, in many cases the total effect of the exposure is of policy interest and conditioning on an intermediate is not necessary.[8,10] Moreover, in some settings, effects conditional on birth weight or gestational age may be of interest but may not necessitate the analytic approaches described here. This will be the case if the exposure or intervention under study in fact occurs after birth (eg, a neonatal intervention).[6] In these cases, birth weight or gestational age becomes a pre-exposure baseline variable, and the approaches we have described here will not be needed. For a neonatal intervention, one could simply condition on birth weight or gestational age without concern about the biases that arise when conditioning on an intermediate. Such analyses simply reduce to ordinary assessment of effect modification in epidemiology. Birth weight may serve as a confounder for a neonatal intervention in addition to potentially being an effect modifier. These settings should be distinguished from those similar to the birthweight paradox in which the variable that one is conditioning on may lie on a pathway from the exposure to the outcome, or is a descendant of variables on the causal pathway.

The approaches we have described in this paper are applicable to perinatal epidemiology more generally. One might, for example, condition on maternal weight gain in assessing the effect of a maternal nutrition supplement on infant mortality. Other examples in the literature include conditioning on preterm birth in assessing the effect of preeclampsia on cerebral palsy[41] or conditioning on infant weight in assessing the effect of a maternal calcium supplement during pregnancy on childhood blood pressure.[42] The approaches we have described here are applicable to these and numerous other settings. As the existing literature has made clear, conditioning on an intermediate in perinatal epidemiology can be problematic and can give rise to severe biases. In many contexts, conditioning on an intermediate is not necessary and is best avoided. Nevertheless, there are cases in which such conditional effects are of scientific or policy interest. We have shown that several alternative approaches can be used to draw inferences in such settings. These methodologic tools are imperfect, make assumptions, and need to be interpreted carefully but they can nonetheless be useful in reasoning about direct and conditional effects.

ACKNOWLEDGMENTS

We thank the reviewers for helpful comments.

REFERENCES

1. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*. 1992;3:143–155.
2. Cole SR, Hernán MA. Fallibility in estimating direct effects. *Int J Epidemiol*. 2002;31:163–165.
3. VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology*. 2010;21:540–551.
4. Kiely JL. Some conceptual problems in multivariable analyses of perinatal mortality. *Paediatr Perinat Epidemiol*. 1991;5:243–257.
5. Kiely JL, Kleinman JC. Birth-weight-adjusted infant mortality in evaluations of perinatal care: towards a useful summary measure. *Stat Med*. 1993;12:377–392.
6. Kramer MS. Biology vs. methodology in investigating causal pathways for infant mortality. *Paediatr Perinat Epidemiol*. 2009;23:414–416.
7. Hernández-Díaz S, Schisterman EF, Hernán MA. The birth-weight "paradox" uncovered? *Am J Epidemiol*. 2006;164:1115–1120.
8. Schisterman EF, Whitcomb BW, Mumford SL, Platt RW. Z-scores and the birthweight paradox. *Paediatr Perinat Epidemiol*. 2009;23:403–413.
9. Whitcomb BW, Schisterman EF, Perkins NJ, Platt RW. Quantification of collider-stratification bias and the birthweight paradox. *Paediatr Perinat Epidemiol*. 2009;23:394–402.
10. Wilcox A, Weinberg CR, Basso O. On the pitfalls of adjusting for gestational age at birth. *Am J Epidemiol.* In press.
11. Judd CM, Kenny DA. Process analysis: estimating mediation in treatment evaluations. *Eval Rev*. 1981;5:602–619.
12. Yerushalmy J. The relationship of parents' cigarette smoking to outcome of pregnancy—implications as to the problem of inferring causation from observed associations. *Am J Epidemiol*. 1971;93:443–456.
13. Wilcox AJ. Birthweight and perinatal mortality: the effect of maternal smoking. *Am J Epidemiol*. 1993;137:1098–1104.
14. Platt RW, Joseph KS, Ananth CV, Grondines J, Abrahamowicz M, Kramer MS. A proportional hazards model with time-dependent covariates and time-varying effects for analysis of fetal and infant death. *Am J Epidemiol*. 2004;160:199–206.
15. Hernán MA, Hernández-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615–625.
16. VanderWeele TJ, Robins JM. Directed acyclic graphs, sufficient causes and the properties of conditioning on a common effect. *Am J Epidemiol*. 2007;166:1096–1104.
17. Glymour MM, Greenland S. Causal diagrams. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia: Lippincott Williams and Wilkins; 2008:chap 12.
18. Cole SR, Platt RW, Schisterman EF, et al. Illustrating bias due to conditioning on a collider. *Int J Epidemiol*. 2010;39:417–420.
19. Ananth CV, Wilcox AJ. Placental abruption and perinatal mortality in the United States. *Am J Epidemiol*. 2001;153:332–337.
20. Basso O, Wilcox AJ. Intersecting birth weight-specific mortality curves: solving the riddle. *Am J Epidemiol*. 2009;169:787–797.

21. Rubin DB. Formal modes of statistical inference for causal effects. *J Stat Plan Inference*. 1990;25:279–292.
22. Hernán MA. A definition of causal effect for epidemiological studies. *J Epidemiol Community Health*. 2004;58:265–271.
23. Petrini J, Damus K, Russell R, Poschman K, Davidoff MJ, Mattison D. Contribution of birth defects to infant mortality in the United States. *Teratology*. 2002;66:S3–S6.
24. Update on overall prevalence of major birth defects—Atlanta, George, 1978–2005. *Morb Mortal Wkly Rep*. 2008;57:1–5.
25. Mili F, Edmonds LD, Khoury MJ, McClearn AB. Prevalence of birth defects among low-birth-weight infants: a population study. *Am J Dis Child*. 1991;145:1313–1318.
26. Pearl J. Direct and indirect effects. In: Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence. San Francisco: Morgan Kaufmann; 2001:411–420.
27. VanderWeele TJ, Hernández-Diaz S. Is there a direct effect of pre-eclampsia on cerebral palsy not through preterm birth? *Paediatr Perinat Epidemiol*. 2011;25:111–115.
28. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics*. 2002;58:21–29.
29. Rubin DB. Direct and indirect effects via potential outcomes. *Scand J Stat*. 2004;31:161–170.
30. Robins JM, Richardson TS, Spirtes P. On identification and inference for direct effects. *Epidemiology*. In press.
31. Sjölander A, Humphreys K, Vansteelandt S, Bellocco R, Palmgren J. Sensitivity analysis for principal stratum direct effects, with an application to a study of physical activity and coronary heart disease. *Biometrics*. 2009;65:514–520.
32. Chiba Y. Bias analysis for the principal stratum direct effect in the presence of confounded intermediate variables. *J Biom Biostat*. 2010;1:101.
33. Joffe M, Small D, Hsu C-Y. Defining and estimating intervention effects for groups that will develop an auxiliary outcome. *Stat Sci*. 2007;22:74–97.
34. Gallop R, Small DS, Lin JY, Elliott MR, Joffe M, Ten Have TR. Mediation analysis with principal stratification. *Stat Med*. 2009;28:1108–1130.
35. Peterson ML, Sinisi SE, van der Laan MJ. Estimation of direct causal effects. *Epidemiology*. 2006;17:276–284.
36. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*. 2009;20:18–26.
37. Robins JM, Richardson TS. Alternative graphical causal models and the identification of direct effects. In: Shrout P, ed. *Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures*. New York: Oxford University Press; 2011.
38. VanderWeele TJ, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Stat Interface*. 2009;2:457–468.
39. VanderWeele TJ, Vansteelandt S. Odds ratios for mediation analysis with a dichotomous outcome. *Am J Epidemiol*. 2010;172:1339–1348.
40. Ananth CV, VanderWeele TJ. Placental abruption and perinatal mortality: impact of preterm delivery assessed through a causal analysis. *Am J Epidemiol*. 2011;174:99–108.
41. Mann J, McDermott S, Griffith M, Hardin J, Gregg A. Uncovering the complex relationship between pre-eclampsia, preterm birth and cerebral palsy. *Paediatr Perinat Epidemiol*. 2011;25:100–110.
42. Belizán JM, Villar J, Bergel E, et al. Long term effect of calcium supplementation during pregnancy on the blood pressure of offspring: follow up of a randomised controlled trial. *Brit Med J*. 1997;315:281–285.