REPRODUCTIVE AND PERINATAL EPIDEMIOLOGY (R PLATT, SECTION EDITOR)

CrossMark

The Curse of the Perinatal Epidemiologist: Inferring Causation Amidst Selection

Jonathan M. Snowden[1,2] · Marit L. Bovbjerg[3] · Mekhala Dissanayake[2] · Olga Basso[4,5]

Abstract

Purpose of Review Human reproduction is a common process and one that unfolds over a relatively short time, but pregnancy and birth processes are challenging to study. Selection occurs at every step of this process (e.g., infertility, early pregnancy loss, and stillbirth), adding substantial bias to estimated exposure-outcome associations. Here, we focus on selection in perinatal epidemiology, specifically, how it affects research question formulation, feasible study designs, and interpretation of results.

Recent Findings Approaches have recently been proposed to address selection issues in perinatal epidemiology. One such approach is the ongoing pregnancies denominator for gestation-stratified analyses of infant outcomes. Similarly, bias resulting from left truncation has recently been termed "live birth bias," and a proposed solution is to control for common causes of selection variables (e.g., fecundity, fetal loss) and birth outcomes. However, these approaches have theoretical shortcomings, conflicting with the foundational epidemiologic concept of populations at risk for a given outcome.

Summary We engage with epidemiologic theory and employ thought experiments to demonstrate the problems of using denominators that include units not "at risk" of the outcome. Fundamental (and commonsense) concerns of outcome definition and analysis (e.g., ensuring that all study participants are at risk for the outcome) should take precedence in formulating questions and analysis approaches, as should choosing questions that stakeholders care about. Selection and resulting biases in human reproductive processes complicate estimation of unbiased causal exposure-outcome associations, but we should not focus solely (or even mostly) on minimizing such biases.

Keywords Selection bias · Population at risk · Birth outcomes · Infertility · Early pregnancy loss · Causation

✉ Jonathan M. Snowden
  snowden@ohsu.edu

  Marit L. Bovbjerg
  marit.bovbjerg@oregonstate.edu

  Mekhala Dissanayake
  dissanam@ohsu.edu

  Olga Basso
  olga.basso@mcgill.ca

1   School of Public Health, Oregon Health and Science University-Portland State University, 3181 SW Sam Jackson Park Rd, Mail Code: CB-669, Portland, OR 97239-3098, USA

2   Department of Obstetrics and Gynecology, Oregon Health and Science University, 3181 SW Sam Jackson Park Rd, Mail Code: L-466, Portland, OR 97239-3098, USA

3   College of Public Health and Human Sciences, Oregon State University, Milam Hall 103, Corvallis, OR 97331, USA

4   Department of Obstetrics & Gynecology, Research Institute of the McGill University Health Centre, Montreal, Canada

5   Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Purvis Hall, 1020 Pine Avenue West Mail Code: QC H4A 3J1, Montreal, QC H3A 1A2, Canada

Introduction: Selection Processes and Risk in Perinatal Epidemiology

In perinatal epidemiology, we seek to establish the effects of exposures on outcomes among dynamic and complex populations: the population of people who may conceive, pregnant women, their fetuses, neonates, infants, and women in the postpartum and inter-conception periods (who may or may not become pregnant again). Some of these populations are not well-defined (e.g., people who may not desire children but are nonetheless at risk for pregnancy), some are difficult to enumerate (e.g., blastocysts and embryos in early gestation), and most of them present challenges to researchers.

These populations are biologically interrelated, and particularly so during pregnancy. Most populations in reproductive and perinatal epidemiology are characterized by key transitions that remove people or gestations from the "at risk" population, altering the pool among whom outcomes can be studied. Beginning before conception, there is an extended process of cohort attrition from embryonic development through birth and early childhood [1–3, 4•]. By the time a woman realizes she is pregnant, the most extensive cohort attrition has already occurred [4•, 5]. We cannot measure the instances of fertilization and can only measure implantations and subsequent losses with great difficulty [4•, 6, 7]. Of the former, it is estimated that only about one third makes it to live birth, be it preterm or term birth [4•, 5]. These processes of selection (e.g., implantations to clinically recognized pregnancies, to birth) and attrition (e.g., the loss of preterm live births and stillbirths that are absent from the population of term live births) determine the populations that we study in perinatal epidemiology (illustrated in Fig. 1 over the course of gestation). An important implication of these transitions is the recognition that the population of live births has been culled significantly by the time we study it (e.g., to examine infant outcomes). Other populations within which we seek to analyze causal effects are similarly affected by selection

processes, e.g., the population of preterm live births, who represent a small and highly selected subset of the gestations that reach viability [8••].

The selection occurring at each of these stages can result in biases of different types, depending on the research question. We employ causal diagrams to illustrate these biases [9, 10]. In studies seeking to establish the causal effect of prematurity on neonatal death [11], confounders include maternal and pregnancy characteristics (e.g., race/ethnicity, multi-fetal gestation) as well as pathologies like preeclampsia, chorioamnionitis, and those that we do not know (Supplemental Fig. S1, Panel A). However, if we restrict our population to preterm births and seek to establish the causal effect of a preterm birth precursor (e.g., preeclampsia) on neonatal death, then the same causal structure results in selection bias (Supplemental Fig. S1, Panel B) [8••, 12]. One cannot simply alter the sampling frame to circumvent this selection bias, as in the classical Berkson bias [13, 14], which can be avoided by recruiting study participants in a setting other than the hospital. For this question, no sample of preterm infants will be immune from this type of selection—except for the hypothetical case in which babies were randomly assigned to being delivered preterm. The forces that determine this sample composition are diverse and include a host of social, environmental, and biological processes that are largely out of the investigator's control [11, 15, 16].

Thus, the selection processes that determine the populations we study in perinatal epidemiology are more fundamental—and more intractable—than in other fields of epidemiology. They are both widespread and largely unobservable. They affect the questions we ask (by defining the scope of causal effects we can hope to estimate), how we define our target population, the study designs we can use, the biases incurred in analyses among these populations, and the methods we may use to control for these biases. In this paper, we discuss these selection processes and their implications for perinatal epidemiology.

Fig. 1 Selection and attrition processes during the course of human reproduction and gestation
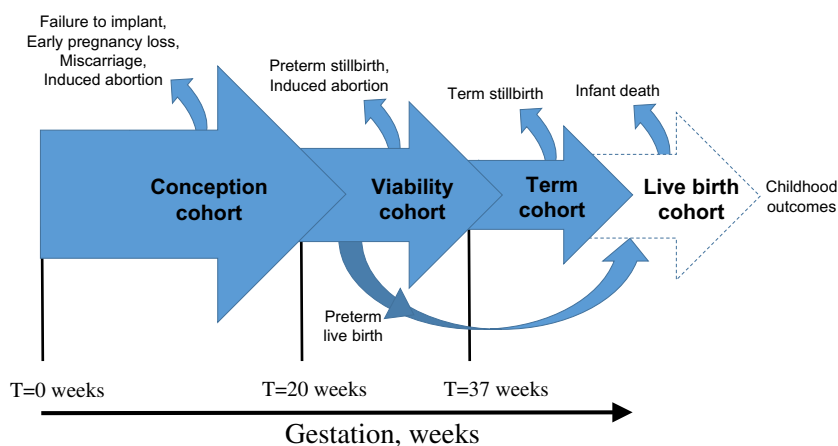
Methodological Considerations for Causal Inference in Perinatal Epidemiology: Recent Developments and Foundational Concepts

As epidemiologic methods have evolved in the last two decades to include newer approaches for estimating causal effects [9, 17–23], there has been a growing focus on applying tools and methods, such as causal diagrams and bias analysis, to address selection issues in perinatal epidemiology [8••, 24••, 25•, 26]. One well-known bias that these advanced tools are being applied to is the bias that affects gestation-stratified analyses of prenatal exposures and postnatal endpoints (resulting from conditioning on gestational-age, a mediator and a collider) [8••, 24••, 27]. The counter-intuitive findings that can result are well-documented (e.g., a seemingly protective effect of preeclampsia on cerebral palsy, infant death, and many other infant morbidities [28–33]), as is their non-causal basis [8••, 24••, 27].

One approach to addressing this particular selection issue is to analyze all pregnancies reaching a given gestation as the population at risk for pregnancy outcomes (variously described as the "ongoing pregnancies denominator" and the "fetuses-at-risk denominator") [34, 35]. This approach prevents conditioning on gestational age after the beginning of the time at risk and has been applied to an increasing number of fetal, maternal, and infant outcomes [36, 37•, 38–40]. This approach is non-controversial for outcomes that occur before the onset of labor (e.g., induction of labor, antepartum stillbirth) [34, 41–42, 43••, 44–45], and is also intuitive: The population at risk for an antepartum stillbirth at or after a given gestation (say, 37 weeks) is not deliveries occurring at 37-weeks' gestation; gestations that continue to 38 weeks, 39 weeks, and beyond were also at risk for this outcome at 37 weeks. Therefore, this approach uses the population of gestations reaching 37 weeks as the denominator for this outcome, regardless of whether birth occurred at 37 weeks or later. Extension of this formulation to neonatal and childhood conditions likely to have a prenatal origin has been proposed [38–40, 46]. However, the application of this approach to postnatal outcomes is controversial and has been shown to result in misleading estimates [43••, 47].

Although it prevents conditioning on gestational age after the beginning of time at risk, use of the ongoing pregnancies denominator for postnatal outcomes promptly runs into another, related problem: It results in the inclusion of denominator units that have not yet reached the beginning of the time at risk [43••, 44]. Studying the role of pre-conception and prenatal factors in relation to endpoints that can only be diagnosed among those having reached a specific milestone (such as live birth, or a given childhood age) represents a common challenge in perinatal epidemiology. In the example above of term antepartum stillbirth risk (i.e., at 37 weeks' gestation or later), the infants who are born preterm (whether liveborn or stillborn) are not counted, nor are gestations that ended in early pregnancy loss. This concern about populations missing due to not reaching the beginning of the time at risk (i.e., left truncation [48, 49]) is also increasingly being discussed (and, when applied to neonatal outcomes occurring after live birth, has been termed "live birth bias"[25•, 26]).

Say that researchers are analyzing the effect of a preconception environmental exposure on the risk of a childhood outcome like autism spectrum disorders (ASD). Suppose that the exposure, like other prenatal and preconception causes of ASD (e.g., advanced parental age, ambient air pollution) [50–52], also increases risk of miscarriage and stillbirth [53–55]. The outcome of ASD, by definition, may only occur and be measured in conceptions that survive to viability and which result in a live birth that reaches a given age (approximately 3 years old, for ASD). Again, there is selection at several steps in the reproductive process (Supplemental Fig. S2). Given that both our environmental exposure of interest and other factors (referred to generally as "Exposure B" in Fig. S2, Panel A) affect the risk of miscarriage, of stillbirth, and of ASD, conditioning on survival beyond each of these steps opens unblocked backdoor paths between exposure and outcome, resulting in bias (Fig. S2, Panel B). One approach that has been proposed is to control for these common causes of the selection variables (e.g., fecundity, fetal loss) and the subsequent outcome (e.g., ASD), in an attempt to close biasing pathways opened up by selection [25•, 56]. However, the meaning of such an estimate, and of the counterfactual it represents, is unclear [47, 57••, 58].

A foundational epidemiologic concept is that of "population at risk" for a given outcome [59, 60••]. Only by rigorously defining, identifying, and sampling the population at risk of the outcome can we validly estimate the effect of an exposure on that outcome. Specifically, all members of the denominator must be able to become members of the numerator (i.e., be at risk of the outcome during follow-up). The following thought experiment shows what can happen when we fail to meet this basic criterion.

A Thought Experiment Demonstrating the Utility of Conditioning on Survival

Let us consider a large double-blind randomized controlled trial to test a treatment, to be initiated before conception, aimed at preventing autism spectrum disorders (ASD). Given that ASD are rare, but have a relatively high risk of recurrence [61, 62], researchers focus on women who have had a first child diagnosed with ASD and are planning a second child. In total, 4000 women are recruited, half of whom are randomized to treatment and half to placebo. Participants attempt conception for up to 12 cycles and collect weekly urine samples to detect implantation. Taking full advantage

of the imaginary nature of this study, we assume full compliance and no dropouts. Even though this is an experimental pre-conception cohort with complete follow-up, providing an estimate based on all randomized women is not necessarily the best option. The treatment may affect any of the steps prior to a child surviving to age 3 (conception, early pregnancy loss, fetal survival, and childhood survival), when the outcome can be measured. The imaginary results of this trial are summarized in Table 1.

The last column shows the relative risk (RR) calculated based on the number of ASD cases divided by the denominator at each step. When all randomized women are considered, as customary in most trials, the treatment appears to reduce the probability of ASD by 36% (i.e., RR = 0.64). Yet, if a pharmaceutical company advertised this figure to demonstrate the effectiveness of their treatment, most would consider it misleading. Half of the drug's apparent effect is due to attrition prior to the time when diagnosis of ASD is even possible: Women in the treatment arm have a 21% lower probability of having a child who survives to age 3, due mostly to decreases in conceptions. Given survival to age 3, the treatment reduces the probability of ASD by 18% (RR = 0.82). This is a much smaller protective effect than the intention-to-treat analysis implied. The overall RR of 0.64, despite being an "unbiased estimate" of the overall treatment effect, is neither very useful nor very transparent. Indeed, if the overall effect were the desired one, an even more impressive result could be achieved by giving women long-acting reversible contraception or sterilization.

It could be argued that the more appropriate estimate is not the relative risk of having a child with ASD but, rather, the relative risk of *having* a child *without* ASD, as that is the desired endpoint. This is expressed by a risk ratio of (710/2000)/(881/2000) = 0.805. From this perspective, the placebo appears to be superior, with a 20% higher probability of

having a child without ASD. However, this estimate answers a different question and obscures the fact that, given that a child survives to age 3, the treatment *does* reduce ASD risk.

This (admittedly artificial) example highlights the difficulty of using a denominator that includes units not at risk, if the exposure differentially affects the probability of reaching the at-risk stage. If (continuing our hypothetical scenario) a brain lesion characteristic of ASD could be detected by ultrasound from week 20 of gestation, estimating the risk of this outcome among all pregnancies surviving to 20 weeks would be appropriate. However, given current clinical capabilities, measuring the risk of actual ASD only among survivors to age 3 is arguably more relevant from a clinical perspective.

Epidemiologists often condition on post-exposure events without agonizing over it. For example, when examining the risk of infertility in women exposed to maternal smoking in utero, not only do we condition on live birth but also on survival to sexual maturity, despite evidence that maternal smoking affects the probability of both these events [63, 64]. Yet, aside from the extreme difficulty of reconstructing *a posteriori* the original pregnancy cohort, it is an incontrovertible (if cynical) fact that infertility is not a concern for those who have died. Furthermore, using the entire cohort would attenuate the effect of prenatal exposure to maternal smoking by including in the study population units that are "prisoners" of the denominator as they cannot experience the outcome.

It is worth noting that, in a study such as the above, the study population is often restricted to pregnancy planners (e.g., [65–67]), which adds a further—and more controversial—layer of conditioning, if children of smokers were less likely to use contraception in a consistent manner (as has been reported for smokers [68]). Unlike the unborn conceptuses and the children who did not reach sexual maturity, those who are not included in a study because they had conceived by accident had a risk of infertility greater than 0, and their

Table 1 Results of a hypothetical randomized controlled trial of a treatment preventing autism spectrum disorders (ASD) among women with a prior child with ASD

	Treatment		Placebo		ASD RR among pregnancies reaching this stage[b]
	N reaching this stage	% reaching this stage[a]	N reaching this stage	% reaching this stage[a]	
All women enrolled	2000		2000		0.64
Chemical pregnancies	1480	74.0	1700	85.0	0.74
Clinical pregnancies	947	64.0	1156	68.0	0.79
Live births	796	84.1	1006	87.0	0.81
Survived to age 3	784	98.5	996	99.0	0.82
ASD	74	9.4	115	11.5	
No ASD	710	90.6	881	88.5	

[a] Percents are conditional on having reached the previous stage

[b] Each relative risk is calculated using as denominators the numbers reported in the column treatment and placebo, for exposed and unexposed, respectively. The numerator is the same in all calculations and is given by 74 and 115

exclusion could lead to overestimating the effect of prenatal smoking on fertility.

While these examples effectively show the potential problems of using denominators that include units not at risk of the outcome, the fact that exposures can differentially affect competing events should, at a minimum, be discussed. This is particularly true when presenting comparisons over time or across populations, as the probability of surviving to given milestones may differ over time or between populations (e.g., survival of very preterm infants).

What Was the Question Again? Incorporating Stakeholder Perspectives into Formulation of Causal Questions

It bears repeating that formulating a good research question is fundamental for sound science. We argue that this principle be applied in perinatal epidemiology, despite the various factors that play into an investigator's scientific and analytical choices (e.g., availability of a given dataset that may drive questions, responsiveness to funding agencies' current priorities, the temptation to use "fashionable" methods). Without discounting the importance of incremental knowledge gains whose public health or clinical applications may only become apparent later, we nonetheless advocate for selecting questions whose results will drive further scientific discovery, policy, and practice (e.g., those that may be translated into public health interventions or policies, and which individual people care about [69, 70]). The preceding thought experiment demonstrates that the stage at which we define our causal question is critical for both estimation and interpretation of the effects. In the population of pre-pregnancy women, preventing conception itself is extremely effective in preventing a case of an adverse childhood outcome. However, non-outcomes are not equal in the eyes of stakeholders: A non-outcome owing to early pregnancy loss does not have the same meaning as a non-outcome owing to a conception that is carried to term, resulting in a child who is ASD-free.

The question "What is the effect of treatment among 3-year-olds?" imposes selection and does not express the full effects of the exposure on all reproductive processes leading up to the outcome. Furthermore, this selection is likely to result in bias because of many other competing risks throughout the reproductive process (Supplemental Fig. S3). There are various stages that a censoring-outcome confounder may introduce bias under selection (Fig. S3, Panel A); however, a censoring-outcome confounder must affect survival to only one such stage to introduce this selection bias (as with Confounder C in Fig. S3, Panel B). Although the full effect of treatment is not captured by this question ("What is the effect of treatment among 3-year-olds?"), and bias due to selection is likely, this is nevertheless a question that people care about. In this case, it is likely

the question that stakeholders care most about, which should be prioritized even as we consider approaches to address selection bias inherent in answering such questions. In these instances, whenever possible, we should also provide estimates of the effect of the exposure on key selection stages (e.g., the probability of clinical pregnancy and live birth).

Choosing a question that is relevant to the people who will use those findings sometimes results in bias, but changing the question to be one that does not incur these biases may make it less relevant. In the thought experiment above, we find the causal question of most interest is, "Compared to placebo, what is the causal effect of the experimental drug on ASD incidence among children at risk for this outcome?" Although this specific outcome and definition of time at risk imposes selection onto the population among whom we estimate the causal effect, these selection factors are motivated by the scientific question. In other words, the censoring processes (illustrated in Figs. 1, S2, and S3) are not nuisance parameters whose consequences we wish to adjust away; rather, they are variables of causal interest that meaningfully impact our scientific question. Specifically, these censoring variables affect how our population at risk is constituted. We see utility in understanding how these censoring processes and competing events may affect our estimated association, but like others [57••], we disagree that the most logical solution is to attempt to adjust away their influence. Doing so is conceptually analogous to redefining the population of interest as all conceptions or all women enrolled in the study. As noted above, this change in focus includes study participants who are never at risk for the outcome and also changes the question to one that patients care less about. Finally, it is not a plausible solution to adjust for all common causes of the outcome (here, ASD) and conception, fetal loss, stillbirth, and infant death. This represents an extremely large number of variables, some of which will likely remain undefined.

Changing the study population changes the causal question, which in turn alters applications of the results—sometimes dramatically. Thinking through, and defining, one's research question in detail is an often overlooked, yet essential, step of the research process. Considerations to address when defining the research question include, who is the target population? How does our sample population differ from this target population? What is the outcome? Who is at risk for it, and over what time? What is the exposure, and how is it temporally related to the outcome? Given exposure and outcome, what are the confounders? How well can we measure each of these variables? What are the likely biases, and how can we adjust for or minimize these in the analysis phase?

In addition to these considerations, we also advocate for addressing the following point in helping guide question formulation: What could be done with these results if our study finds evidence of an association (or if it does not)? All these considerations matter, and the weight given to each depends on the investigator and the specific project. When studying the

effects of prenatal or preconception exposures on childhood endpoints, focusing too much attention on one problem (e.g., how do we address bias owing to selection processes in human reproduction?) risks losing sight of other concerns. We argue that fundamental (and commonsense) concerns of outcome definition and analysis (e.g., ensuring that all study participants are at risk for the outcome) should take precedence, as should choosing questions that stakeholders care about—not just reducing bias.

Conclusion

Selection and resulting biases are omnipresent in human reproductive and perinatal processes. In fact, these processes of transition, attrition, and survival are arguably at the core of human reproduction [4•]. They complicate our task to estimate unbiased exposure-outcome associations, but it may not be in our best interest to focus solely (or even mostly) on minimizing such biases. To illustrate why, we present one last thought experiment, the most extreme so far.

Let us consider the ideal study design to study some adult outcome and contrast this with the optimal study design to study birth outcomes. Say we are interested in whether a given dietary pattern affects risk of developing incident hypertension among adults in their 40s and 50s. We can imagine enrolling a large sample of normotensive adults in their 20s, in a perfect universe with easy recruitment methods, excellent participation rates, and high retention throughout our desired study period. We can imagine a world where participants would gladly comply with whichever dietary regimen they are randomly assigned, from a large and diverse list of possibilities. We could follow these people for years or decades, and track their blood pressure trajectories, hypertension incidence, and a host of other risk factors and outcomes for good measure.

Now, picture conducting an analogous study to estimate the effects of various dietary regimens before and during pregnancy on a childhood outcome (say obesity, or ASD again). Even with the low administrative burdens, easy enrollment and retention, and sample with high compliance, this task is considerably more difficult by comparison. In our hypertension study, we had resources to track the few losses to follow up. The rate of mortality was not very high among our age group, so few participants were censored due to death. In contrast, in our preconception cohort, no matter how well-funded or beautifully-designed our study is, we can count on at least 25% of the original conception population at risk for childhood outcomes being lost, many of them before we can even enumerate them or know of their existence. Worse still, the losses may be differential due to the exposure. It is difficult to draw a parallel to the adult blood pressure study, but perhaps, there is a cataclysmic disaster or an alien invasion that removes a quarter of the population from the adult study, without our

being able to track them (or even, precisely how many were taken). In the adult study, we at least were aware of the existence of all study participants, but not so in our preconception cohort. In preconception cohorts, all women are enumerated and known, but the internal and hidden nature of conception and early embryonic development means that a large share of our potential at-risk population cannot even be detected. Are women not able to get pregnant? Losing conceptuses early? Losing embryos later but still before the pregnancy is recognized? Again, it is difficult to imagine how we could explain this challenge to our colleagues in adult cardiovascular epidemiology—perhaps there is an impenetrable, opaque force-field that prevents some unknown proportion of the participants in their study from being enumerated, or even having their existence known to the study team. It is even harder to devise an analogy for infertility, yet another selection factor that can be affected by our exposure of interest (like early pregnancy loss)—particularly since some infertile couples will never know their status, if they never try to conceive.

There may be more apt analogies to describe the challenges facing perinatal epidemiologists in assessing causal effects amidst the dynamic, hidden, and interconnected populations of pregnancy, birth, and childhood, but we believe that no realistic ones can fully capture the dynamics of the many selection processes we have focused on here. What we propose is that, rather than trying to combat the alien invaders to recapture lost study participants, or to see through opaque and impenetrable force-fields, we acknowledge the challenges facing us and use our epidemiologic tools to understand them as best as we can. Thus, adjusting away selection may not always be possible in perinatal epidemiology, but we should remain vigilant in seeking to understand these processes and how they affect our results. We should also be mindful to formulate analytical approaches that conform to the foundational tenets of epidemiology (e.g., populations at risk), and questions that matter to stakeholders. When this process takes us down the path of confronting bias owing to selection processes and competing risks, we should do our best to understand and address these biases, while still letting our question drive the analytical approach.

Compliance with Ethical Standards

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance
•• Of major importance

1. Wilcox AJ, Weinberg CR, O'Connor JF, Baird DD, Schlatterer JP, Canfield RE, et al. Incidence of early loss of pregnancy. N Engl J Med. 1988;319(4):189–94. https://doi.org/10.1056/NEJM198807283190401.

2. Wilcox AJ, Baird DD, Weinberg CR. Time of implantation of the conceptus and loss of pregnancy. N Engl J Med. 1999;340(23):1796–9. https://doi.org/10.1056/NEJM199906103402304.

3. Zinaman MJ, Clegg ED, Brown CC, O'Connor J, Selevan SG. Estimates of human fertility and pregnancy loss. Fertil Steril. 1996;65(3):503–9.

4.• Macklon NS, Geraedts JP, Fauser BC. Conception to ongoing pregnancy: the 'black box' of early pregnancy loss. Hum Reprod Update. 2002;8(4):333–43 **Provides an accessible introduction to human fertility, conception, and early pregnancy loss.**

5. Chard. T. Frequency of implantation and early pregnancy loss in natural cycles. Baillieres Clin Obstet Gynaecol. 1991;5(1):179–89.

6. Hertig AT, Rock J, Adams EC, Menkin MC. Thirty-four fertilized human ova, good, bad and indifferent, recovered from 210 women of known fertility; a study of biologic wastage in early human pregnancy. Pediatrics. 1959;23(1 Part 2):202–11.

7. Buster JE, Bustillo M, Rodi IA, Cohen SW, Hamilton M, Simon JA, et al. Biologic and morphologic development of donated human ova recovered by nonsurgical uterine lavage. Am J Obstet Gynecol. 1985;153(2):211–7.

8.•• Snowden JM, Basso O. Causal inference in studies of preterm babies: a simulation study. BJO Int J Obstet Gynaecol. 2018;125(6):686–92. https://doi.org/10.1111/1471-0528.14942 **Employs data simulation to demonstrate selection issues as applied to estimating effects of antepartum risk factors on neonatal endpoints in samples restricted to preterm births. Argues that this selection bias is an intractable feature of such analyses, not amenable to analytical correction.**

9. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology. 1999;10(1):37–48.

10. Williams TC, Bach CC, Matthiesen NB, Henriksen TB, Gagliardi L. Directed acyclic graphs: a tool for causal studies in paediatrics. Pediatr Res. 2018. https://doi.org/10.1038/s41390-018-0071-3.

11. Institute of Medicine (IOM). Preterm Birth: Causes, Consequences, and Prevention. In: Behrman RE, Butler AS, editors. The National Academies Collection: Reports. Washington (DC): National Institutes of Health; 2007.

12. Savitz DA. Only some questions of cause and effect can be evaluated in highly selected populations. BJOG Int J Obstet Gynaecol. 2018;125(6):647–8. https://doi.org/10.1111/1471-0528.15001.

13. Berkson J. Limitations of the application of fourfold table analysis to hospital data. Biometrics. 1946;2(3):47–53.

14. Westreich D. Berkson's bias, selection bias, and missing data. Epidemiology. 2012;23(1):159–64. https://doi.org/10.1097/EDE.0b013e31823b6296.

15. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. Lancet. 2008;371(9606):75–84. https://doi.org/10.1016/S0140-6736(08)60074-4.

16. Romero R, Espinoza J, Kusanovic JP, Gotsch F, Hassan S, Erez O, et al. The preterm parturition syndrome. BJOG Int J Obstet Gynaecol. 2006;113 Suppl 3:17–42. https://doi.org/10.1111/j.1471-0528.2006.01120.x.

17. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000;11(5):550–60.

18. Greenland S. Basic methods for sensitivity analysis of biases. Int J Epidemiol. 1996;25(6):1107–16.

19. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. Epidemiology. 2006;17(3):268–75.

20. Hernan MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? Epidemiology. 2006;17(4):360–72. https://doi.org/10.1097/01.ede.0000222409.00878.37.

21. Lash TL, Abrams B, Bodnar LM. Comparison of bias analysis strategies applied to a large data set. Epidemiology. 2014;25(4):576–82. https://doi.org/10.1097/EDE.0000000000000102.

22. Mayeda ER, Tchetgen Tchetgen EJ, Power MC, Weuve J, Jacqmin-Gadda H, Marden JR, et al. A simulation platform for quantifying survival Bias: an application to research on determinants of cognitive decline. Am J Epidemiol. 2016;184(5):378–87. https://doi.org/10.1093/aje/kwv451.

23. Naimi AI, Cole SR, Kennedy EH. An introduction to g methods. Int J Epidemiol. 2017;46(2):756–62. https://doi.org/10.1093/ije/dyw323.

24.•• Wilcox AJ, Weinberg CR, Basso O. On the pitfalls of adjusting for gestational age at birth. Am J Epidemiol. 2011;174(9):1062–8. https://doi.org/10.1093/aje/kwr230 **Employs causal diagrams to demonstrate the bias that results from conditioning on gestational age in analyses of antepartum exposures and neonatal endpoints, drawing a clear distinction between descriptive and etiological research questions.**

25.• Liew Z, Olsen J, Cui X, Ritz B, Arah OA. Bias from conditioning on live birth in pregnancy cohorts: an illustration based on neurodevelopment in children after prenatal exposure to organic pollutants. Int J Epidemiol. 2015;44(1):345–54. https://doi.org/10.1093/ije/dyu249 **Establishes the causal and theoretical basis for "live-birth bias" resulting from conditioning on live-birth in analysis of child outcomes, and proposes the analytical solution of controlling for censoring-outcome confounders in the statistical model, to block biasing pathways between exposure and outcome.**

26. Suarez EA, Landi SN, Conover MM, Jonsson Funk M. Bias from restricting to live births when estimating effects of prescription drug use on pregnancy complications: a simulation. Pharmacoepidemiol Drug Saf. 2018;27(3):307–14. https://doi.org/10.1002/pds.4387.

27. Kramer MS, Zhang X, Platt RW. Analyzing risks of adverse pregnancy outcomes. Am J Epidemiol. 2014;179(3):361–7. https://doi.org/10.1093/aje/kwt285.

28. Murphy DJ, Sellers S, MacKenzie IZ, Yudkin PL, Johnson AM. Case-control study of antenatal and intrapartum risk factors for cerebral palsy in very preterm singleton babies. Lancet. 1995;346(8988):1449–54.

29. Wilson-Costello D. Risk factors for neurologic impairment among very low-birth-weight infants. Semin Pediatr Neurol. 2001;8(2):120–6.

30. Perlman JM, Risser RC, Gee JB. Pregnancy-induced hypertension and reduced intraventricular hemorrhage in preterm infants. Pediatr Neurol. 1997;17(1):29–33.

31. Shankaran S, Bauer CR, Bain R, Wright LL, Zachary J. Prenatal and perinatal risk and protective factors for neonatal intracranial hemorrhage. National Institute of Child Health and Human Development neonatal research network. Arch Pediatr Adolesc Med. 1996;150(5):491–7.

32. Yu XD, Branch DW, Karumanchi SA, Zhang J. Preeclampsia and retinopathy of prematurity in preterm births. Pediatrics. 2012;130(1):e101–7. https://doi.org/10.1542/peds.2011-3881.

33. Fortes Filho JB, Costa MC, Eckert GU, Santos PG, Silveira RC, Procianoy RS. Maternal preeclampsia protects preterm infants against severe retinopathy of prematurity. J Pediatr. 2011;158(3): 372–6. https://doi.org/10.1016/j.jpeds.2010.08.051.

34. Yudkin PL, Wood L, Redman CW. Risk of unexplained stillbirth at different gestational ages. Lancet. 1987;1(8543):1192–4.

35. Smith GC. Life-table analysis of the risk of perinatal death at term and post term in singleton pregnancies. Am J Obstet Gynecol. 2001;184(3):489–96. https://doi.org/10.1067/mob.2001.109735.

36. Caughey AB, Stotland NE, Escobar GJ. What is the best measure of maternal complications of term pregnancy: ongoing pregnancies or pregnancies delivered? Am J Obstet Gynecol. 2003;189(4):1047–52.

37.• Auger N, Gilbert NL, Naimi AI, Kaufman JS. Fetuses-at-risk, to avoid paradoxical associations at early gestational ages: extension to preterm infant mortality. Int J Epidemiol. 2014;43(4):1154–62. https://doi.org/10.1093/ije/dyu011 **Applies the ongoing pregnancies denominator to neonatal death after preterm birth, arguing that this solution resolves paradoxical associations observed among gestation-stratified analyses of prenatal exposures and neonatal endpoints.**

38. Auger N, Naimi AI, Fraser WD, Healy-Profitos J, Luo ZC, Nuyt AM, et al. Three alternative methods to resolve paradoxical associations of exposures before term. Eur J Epidemiol. 2016;31(10): 1011–9. https://doi.org/10.1007/s10654-016-0175-1.

39. Joseph KS. A consilience of inductions supports the extended fetuses-at-risk model. Paediatr Perinat Epidemiol. 2016;30(1):11–7. https://doi.org/10.1111/ppe.12260.

40. Joseph KS, Kramer MS. The fetuses-at-risk approach: survival analysis from a fetal perspective. Acta Obstet Gynecol Scand. 2018;97(4):454–65. https://doi.org/10.1111/aogs.13194.

41. Smith GC. Estimating risks of perinatal death. Am J Obstet Gynecol. 2005;192(1):17–22. https://doi.org/10.1016/j.ajog.2004.08.014.

42. Caughey AB, Nicholson JM, Cheng YW, Lyell DJ, Washington AE. Induction of labor and cesarean delivery by gestational age. Am J Obstet Gynecol. 2006;195(3):700–5.

43.•• Basso O. Implications of using a fetuses-at-risk approach when fetuses are not at risk. Paediatr Perinat Epidemiol. 2016;30(1):3–10. https://doi.org/10.1111/ppe.12254 **Demonstrates the misleading estimates that result from applying an ongoing pregnancies denominator to analysis of postnatal endpoints.**

44. Caughey AB, Snowden JM. Measuring perinatal complications: different approaches depending on who is at risk. Paediatr Perinat Epidemiol. 2016;30(1):23–4. https://doi.org/10.1111/ppe.12257.

45. Smith GC. Quantifying the risk of different types of perinatal death in relation to gestational age: researchers at risk of causing confusion. Paediatr Perinat Epidemiol. 2016;30(1):18–9. https://doi.org/10.1111/ppe.12259.

46. Joseph KS. Incidence-based measures of birth, growth restriction, and death can free perinatal epidemiology from erroneous concepts of risk. J Clin Epidemiol. 2004;57(9):889–97. https://doi.org/10.1016/j.jclinepi.2003.11.018.

47. Basso O. No rates were harmed in the making of this paper: response to critiques. Paediatr Perinat Epidemiol. 2016;30(1):25–7. https://doi.org/10.1111/ppe.12266.

48. Cain KC, Harlow SD, Little RJ, Nan B, Yosef M, Taffe JR, et al. Bias due to left truncation and left censoring in longitudinal studies of developmental and disease processes. Am J Epidemiol. 2011;173(9):1078–84. https://doi.org/10.1093/aje/kwq481.

49. Howards PP, Hertz-Picciotto I, Poole C. Conditions for bias from differential left truncation. Am J Epidemiol. 2007;165(4):444–52.

50. Gardener H, Spiegelman D, Buka SL. Prenatal risk factors for autism: comprehensive meta-analysis. Br J Psychiatry. 2009;195(1): 7–14. https://doi.org/10.1192/bjp.bp.108.051672.

51. Nybo Andersen AM, Wohlfahrt J, Christens P, Olsen J, Melbye M. Maternal age and fetal loss: population based register linkage study. BMJ. 2000;320(7251):1708–12.

52. Lyall K, Schmidt RJ, Hertz-Picciotto I. Maternal lifestyle and environmental risk factors for autism spectrum disorders. Int J Epidemiol. 2014;43(2):443–64. https://doi.org/10.1093/ije/dyt282.

53. de la Rochebrochard E, Thonneau P. Paternal age and maternal age are risk factors for miscarriage; results of a multicentre European study. Hum Reprod. 2002;17(6):1649–56.

54. Faiz AS, Rhoads GG, Demissie K, Kruse L, Lin Y, Rich DQ. Ambient air pollution and the risk of stillbirth. Am J Epidemiol. 2012;176(4):308–16. https://doi.org/10.1093/aje/kws029.

55. Checa Vizcaino MA, Gonzalez-Comadran M, Jacquemin B. Outdoor air pollution and human infertility: a systematic review. Fertil Steril. 2016;106(4):897–904 e1. https://doi.org/10.1016/j.fertnstert.2016.07.1110.

56. Liew Z, Olsen J, Cui X, Ritz B, Arah OA. Response to Werler and Parker letter: comment on live-birth bias in pregnancy cohorts. Int J Epidemiol. 2015;44(3):1080–1. https://doi.org/10.1093/ije/dyv140.

57.•• Werler MM, Parker SE. Bias from conditioning on live-births in pregnancy cohorts: an illustration based on neurodevelopment in children after prenatal exposure to organic pollutants (Liew et al. 2015). Int J Epidemiol. 2015;44(3):1079–80. https://doi.org/10.1093/ije/dyv139 **Responding to the proposal to adjust away "live birth bias" [proposed in reference 25], argues that it is illogical to adjust away competing risks on fetal death to estimate associations between preconception exposures and child outcomes. Points out the fallacy of including units not at risk for the childhood outcome (e.g., fetuses who did not survive to infancy) in the denominator.**

58. Jones HE, Schooling CM. Let's require the "T-word". Am J Public Health. 2018;108(5):624. https://doi.org/10.2105/AJPH.2018.304365.

59. Greenland S, Rothman KJ. Chapter 3: measure of occurrence. In: Rothman KJ, Greenland S, Lash TL, editors. Modern Epidemiology. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008.

60.•• Szklo M, Nieto FJ. Chapter 2: Measuring Disease Occurrence. Epidemiology: Beyond the Basics. 3rd ed. Burlington: Jones & Bartlett Learning; 2012. **Provides clear substantiation of the fallacy of including in the denominator units not at risk for the outcome, stating that incidence "…is represented by the number of events occurring in a defined population over a specified period of time (numerator), divided by the population at risk for that event over that time (denominator)" [page 49].**

61. Gronborg TK, Schendel DE, Parner ET. Recurrence of autism spectrum disorders in full- and half-siblings and trends over time: a population-based cohort study. JAMA Pediatr. 2013;167(10):947–53. https://doi.org/10.1001/jamapediatrics.2013.2259.

62. Ozonoff S, Young GS, Carter A, Messinger D, Yirmiya N, Zwaigenbaum L, et al. Recurrence risk for autism spectrum disorders: a baby siblings research consortium study. Pediatrics. 2011;128(3):e488–95. https://doi.org/10.1542/peds.2010-2825.

63. Pineles BL, Park E, Samet JM. Systematic review and meta-analysis of miscarriage and maternal exposure to tobacco smoke during pregnancy. Am J Epidemiol. 2014;179(7):807–23. https://doi.org/10.1093/aje/kwt334.

64. Hofvendahl EA. Smoking in pregnancy as a risk factor for long-term mortality in the offspring. Paediatr Perinat Epidemiol. 1995;9(4):381–90.

65. Jensen TK, Henriksen TB, Hjollund NH, Scheike T, Kolstad H, Giwercman A, et al. Adult and prenatal exposures to tobacco smoke as risk indicators of fertility among 430 Danish couples. Am J Epidemiol. 1998;148(10):992–7.

66. Weinberg CR, Wilcox AJ, Baird DD. Reduced fecundability in women with prenatal exposure to cigarette smoking. Am J Epidemiol. 1989;129(5):1072–8.

67. Ye X, Skjaerven R, Basso O, Baird DD, Eggesbo M, Cupul Uicab LA, et al. In utero exposure to tobacco smoke and subsequent reduced fertility in females. Hum Reprod. 2010;25(11):2901–6. https://doi.org/10.1093/humrep/deq235.

68. Baird DD, Weinberg CR, Schwingl P, Wilcox AJ. Selection bias associated with contraceptive practice in time-to-pregnancy studies. Ann N Y Acad Sci. 1994;709:156–64.

69. Fleurence R, Selby JV, Odom-Walker K, Hunt G, Meltzer D, Slutsky JR, et al. How the Patient-Centered Outcomes Research Institute is engaging patients and others in shaping its research agenda. Health Aff. 2013;32(2):393–400. https://doi.org/10.1377/hlthaff.2012.1176.

70. Fleurence RL, Forsythe LP, Lauer M, Rotter J, Ioannidis JP, Beal A, et al. Engaging patients and stakeholders in research proposal review: the patient-centered outcomes research institute. Ann Intern Med. 2014;161(2):122–30. https://doi.org/10.7326/M13-2412.