

REVIEW ARTICLE OPEN



Reevaluating the protective effect of smoking on preeclampsia risk through the lens of bias

Merida Rodriguez-Lopez^{1,2}, Maria Fernanda Escobar^{1,3,4}, Juan Merlo⁵ and Jay S. Kaufman⁵

© The Author(s) 2023

Preeclampsia is a hypertensive disorder that is usually diagnosed after 20 weeks' gestation. Despite the deleterious effect of smoking on cardiovascular disease, it has been frequently reported that smoking has a protective effect on preeclampsia risk and biological explanations have been proposed. However, in this manuscript, we present multiple sources of bias that could explain this association. First, key concepts in epidemiology are reviewed: confounder, collider, and mediator. Then, we describe how eligibility criteria, losses of women potentially at risk, misclassification, or performing incorrect adjustments can create bias. We provide examples to show that strategies to control for confounders may fail when they are applied to variables that are not confounders. Finally, we outline potential approaches to manage this controversial effect. We conclude that there is probably no single epidemiological explanation for this counterintuitive association.

Journal of Human Hypertension (2023) 37:338–344; <https://doi.org/10.1038/s41371-023-00827-9>

THE PUZZLING ASSOCIATION AND ITS BIOLOGICAL EXPLANATION

Preeclampsia (PE) is conceptualized as a hypertensive disorder that affects human pregnancy after 20 weeks of gestational age (GA) [1]. It is a public health problem that increases the risk of maternal morbidity and mortality and relates to long-term cardiovascular diseases [2]. Most cases become clinically evident during the third trimester [3]. The cumulative incidence of PE increases from 0.44% between the 20 and 25th week to 3.8% in the 41st week of gestation [3]. Proteinuria together with a rise in blood pressure (BP) and signs/symptoms of organ and uteroplacental dysfunction are features of the disease's clinical definition. More recently, PE is diagnosed even without proteinuria [4]. For many years this multisystem disorder was thought to be caused by placental insufficiency; however, recent studies have demonstrated that cardiac remodeling can be present even before the onset of PE [5], suggesting a cardiovascular origin for PE [6].

Many observational [7–10] and systematic reviews [11–13] have pointed to a protective effect of smoking in PE risk. This puzzling association has been reported for gestational hypertension (GH) (OR 0.74, 95% CI 0.69–0.79) and PE (OR 0.65, 95% CI 0.58–0.73) [13], for term or preterm PE, for smoking before [14] or during pregnancy [15], and for cohort (RR 0.68, 95% CI 0.67–0.69) and case-control studies (OR 0.68, 95% CI 0.57–0.81) [16]. In contrast, other studies have shown that smoking during pregnancy increases the risk of PE [17], increases the risk just among taller women [18] or increases both systolic and diastolic BP with no association with PE [19]. This protective effect of tobacco on PE risk contrasts with its deleterious effect on severe hypertension and cardiovascular diseases in the population [20, 21]. PE and chronic hypertension (CH), share similar risk factor profiles except

for smoking. For these reasons, this association is still a source of debate.

Arguments in favor of the causal effect of smoking on PE highlight the role of nitric oxide, a potent vasodilator [22]. However, smokers have decreased levels of circulating nitric oxide metabolites [12]. It has been proposed that combustion products of tobacco, including carbon monoxide but not constituents of tobacco, such as nicotine, are responsible for the protective effect [15]. Carbon monoxide mimics nitric oxide and therefore might replicate its effects, inhibiting placental apoptosis, necrosis, and the production of antiangiogenic proteins such as sFlt1 [22]. However, higher levels of blood carboxyhemoglobin, a stable complex of carbon monoxide and hemoglobin, are associated with a higher risk of PE. Animal models have also pointed to an anti-inflammatory effect of nicotine, which might attenuate its capacity to increase systemic BP [23].

Unexpected associations between smoking and health outcomes have been previously reported in the field of perinatal epidemiology [24, 25]. For example, smoking is a risk factor for low birth weight and mortality. However, among those born with low weight, smoking appears to work as a protective factor for mortality [24]. These apparent paradoxes have been explained by selection bias [26]. Recently, some papers have addressed the protective effect of smoking on PE considering possible sources of selection bias one at a time [27, 28]. In this manuscript we suggest multiple mechanisms for bias without going deeper into their statistical structure. We first review key concepts in epidemiology: confounders, mediators, and colliders. Then we discuss the potential impact of selection bias due to exclusion, losses of women and pregnancies, misclassification, and covariate over-adjustment [29]. Finally, we outline some potential strategies to better understand this puzzling association.

¹Faculty of Health Science, Universidad Icesi, Cali, Colombia. ²Unit for Social Epidemiology, Faculty of Medicine, Lund University, Malmö, Sweden. ³Clinical Research Center, Fundación Valle del Lili, Cali, Colombia. ⁴Department of Obstetrics and Gynecology, Fundación Valle del Lili, Cali, Colombia. ⁵Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montréal, QC, Canada. ✉email: mrodriguez@icesi.edu.co

Received: 9 February 2023 Revised: 18 March 2023 Accepted: 27 March 2023

Published online: 11 April 2023

KEY DEFINITIONS IN EPIDEMIOLOGY

Most readers are familiar with the definition of a confounder; however, the concepts of colliders and mediators may be less frequently used among clinical researchers. A “confounder” is commonly understood as a variable that meets three conditions: (1) it is an independent risk factor for the outcome; (2) it is related to the exposure without being affected by the exposure and consequently, (3) it does not lie on the pathway between exposure and the outcome [30]. More generally, a confounder is a common cause of the exposure and the outcome, or a proxy for such a cause [31]. The definition of confounding in a particular exposure-outcome association relies on the structure of the relationships between variables in each causal framework [32]. Well-known strategies to control confounding include but are not limited to restriction, matching, stratification, adjustment, propensity scores, and randomization. In some cases, a variable appears to meet the three conditions listed above but does not generate confounding because, rather than being a common cause, it is a common effect of the exposure and the outcome.

One specific common effect is known as a “collider” [26]. In general, a collider is a third variable that is influenced by both the exposure and the outcome (or by a cause of the exposure and a cause of the outcome). For example, PE and smoking can both lead to renal diseases, then renal disease is a common effect of both PE and smoking. Contrary to confounders, if the collider is controlled for by design or analysis, it can induce a spurious association between the exposure and the outcome which is known as collider bias [33]. A more detailed definition of collider can be found elsewhere [34, 35]. Moreover, renal disease can cause PE, but, as it cannot cause smoking, then it is not a confounder unless a diagnosis influences a smoker to quit. In some causal frameworks, smoking could cause renal disease and then renal disease could cause PE. If this were the case, renal disease can be considered a “mediator” of the effect of smoking on PE. A mediator is a consequence of the exposure and a subsequent cause of the outcome. That is, a variable that lies on the pathway between exposure and outcome. Strategies to reduce confounding bias might also fail if they are applied to mediators without additional assumptions, as it is discussed below.

Figure 1 illustrates the roles of variables and the ways in which they are causally connected [26, 36]. One graphical representation of these roles is known as a directed acyclic graph (DAG), which is a valuable tool for causal inference [31]. In DAGs, the distinction between confounders, mediators, and colliders is made explicit, such as (1) we might want to separate the direct and indirect effects (the effects through the mediator) of an exposure, and (2) controlling for confounders can reduce bias, but controlling for a collider can increase bias [36]. In terms of formal DAG rules,

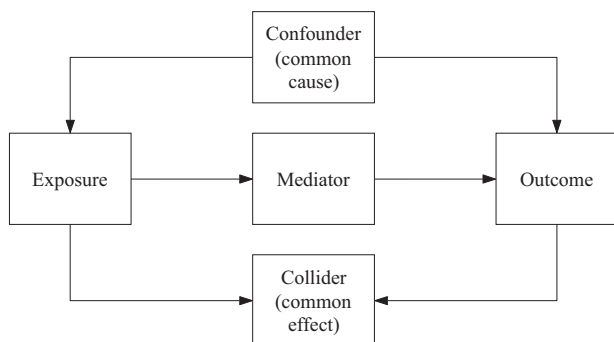


Fig. 1 Basic structure of the relationships between variables within a causal framework perspective. Arrows indicate the direction of causality. Confounder, mediator, and collider refer to the exposure-outcome association.

stratification and conditioning on a confounder block but performing such strategies on a collider unblocks a “backdoor path” between the exposure and the outcome [26]. Up to date, there are other statistical approaches based on standardization (weighting), g-methods and doubly robust methods [37], rather than conditioning, that enable accounting for colliders without generating collider stratification bias.

THE EPIDEMIOLOGICAL VIEW OF THE PUZZLING ASSOCIATION

Eligibility criteria and losses to follow-up of at-risk women

PE is generally diagnosed after 20 weeks of GA. However, there are rare PE cases diagnosed before 20 weeks, mostly associated with molar pregnancy, triploid or antiphospholipid syndrome. Recently, some PE cases have been identified even in the absence of such pathologies [38]. A high BP value <20 weeks is most likely diagnosed as CH [39]. Consequently, the follow-up of women usually starts at 20 weeks of GA in cohort studies, and when using prevalent cases as in case-control and birth-registered studies, those women <20 weeks are excluded from the analysis [10, 40, 41]. Under the current clinical and epidemiological definition of PE, it can only present itself among those pregnancies that continue beyond 20 weeks of gestation. Therefore, there is a time frame within which, even if there is smoke exposure, PE cannot be diagnosed. The restriction of the study population to those pregnancies >20 weeks create a truncated cohort as depicted in Fig. 2 (left side) that can generate bias if smoking is associated with the cause of losses of at-risk women.

Pre-pregnancy smoking doubles the risk of early pregnancy loss [42, 43], including miscarriage [44], with most cases occurring before 14 weeks [45]. Overall, 10–20% of clinically recognized pregnancies will end in early pregnancy loss [46]. The true incidence of miscarriage may increase to 30% [47], as many miscarriages occur before a mother realizes she is pregnant. In fact, the incidence ranges from 18 to 62% when using human chorionic gonadotropin for pregnancy detection [48]. It has also been reported that up to 78% of fertilized ova may be lost. Therefore, the number of conceptions is much higher than the number of detected pregnancies, which is also higher than the number of deliveries. Early pregnancy loss produces a subpopulation cohort that is depleted of those susceptible to the smoking effect at the time by which PE is diagnosed, which might create selection bias at baseline [49]. This phenomenon was demonstrated using a simulation study by Lisonkova and Joseph [27]; however, it did not fully explain the inverse and counterintuitive findings [50].

Competing events

Smoking also increases the risk of adverse perinatal outcomes during the second and third trimesters, which also prevents the incidence of PE by ending the pregnancy. Smoking during pregnancy has been associated with stillbirths after 22 weeks [51], and a doubling of the risk of a premature birth before 34 weeks [52], placental abruption, and fetal growth restriction. These and other conditions lead to an increased likelihood of induced labors and elective cesarean delivery [53]. They are then considered competing events, as shown in Fig. 2 (right-hand side), because their occurrence precludes the onset of PE. During the follow-up of a pregnant woman, the end of pregnancy can be due to a healthy birth, a birth due to PE, or a birth resulting from another complication such as fetal distress, fetal growth restriction, or placental abruption. These complications before the onset of PE impede its occurrence because they usually lead to medical intervention and iatrogenic preterm births [53]. They causally ‘compete’ with PE for the pregnancy’s end.

The competing event could be considered a collider: it is caused by smoking and share common (and often unmeasured) causes

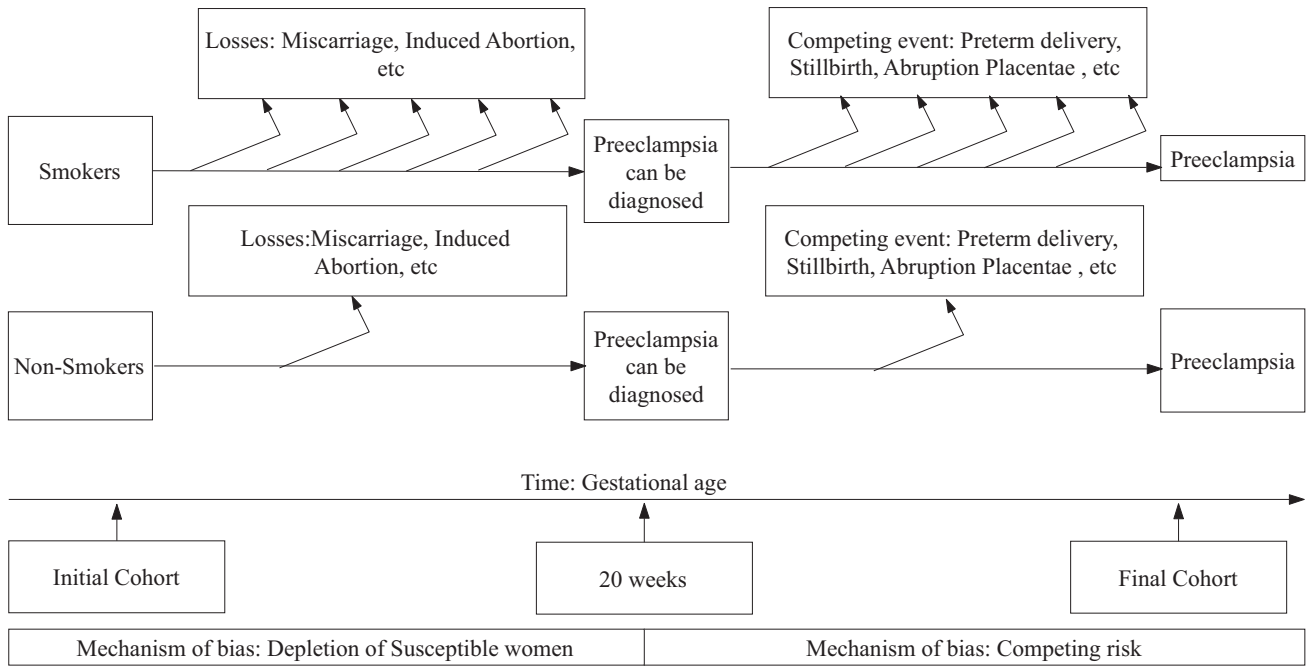


Fig. 2 Potential losses of women at risk during pregnancy. The number of arrows moving out from horizontal lines represent the loss of follow-up among smokers and non-smokers. Left- (below 20 weeks of gestational age) and right-hand sides (above 20 weeks of gestational age) draw two potential mechanisms of bias: the left side shows that losses before 20 weeks create a study cohort that is depleted of cases susceptible to the effect of smoking, while the right side shows that competing events, which are known to be more frequent among smokers, preclude the incidence of PE in that subgroup.

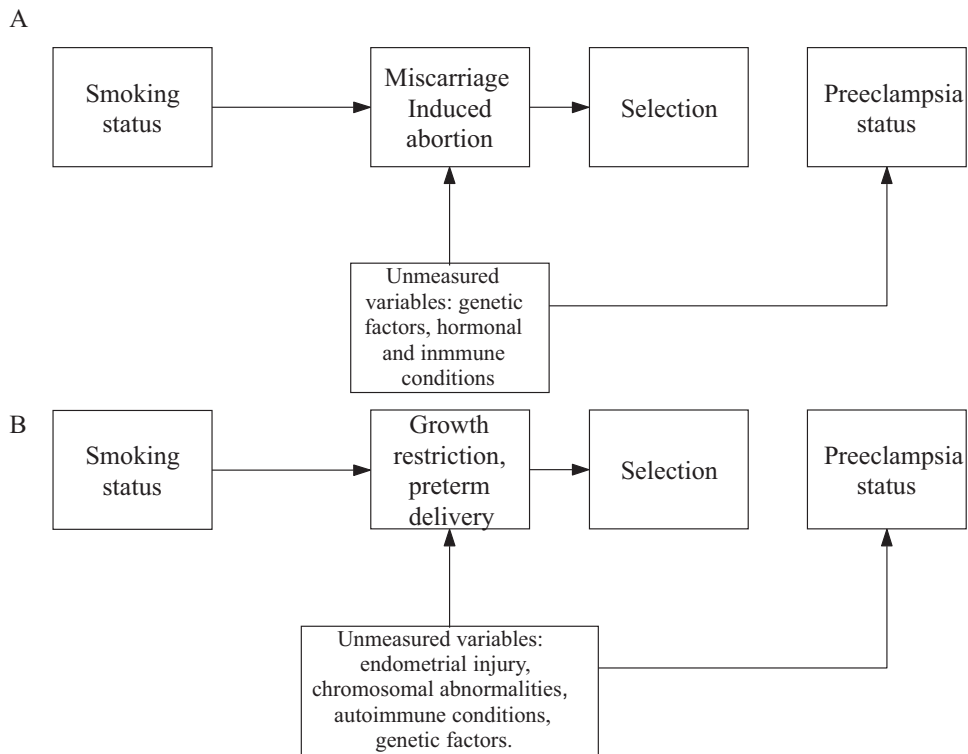


Fig. 3 Collider bias occurs when the selection of the study population is restricted to pregnancies which survive the effect of tobacco. **A** During the first trimester **(B)** During the third Trimester. Arrows indicate the direction of causality. Gray line represents the association of interest.

with PE. Figure 3 represents potential unmeasured confounders in the relation between the collider and PE before (A) and after (B) week 20. If it is not possible to control such relevant unmeasured factors, they collide with smoking in the competing event, and the

path between smoking and PE is open according to DAG rules. If unmeasured variables were measured, they could be included in the statistical model as confounders of the relation between the collider and the outcome, and then the bias could, in principle, be

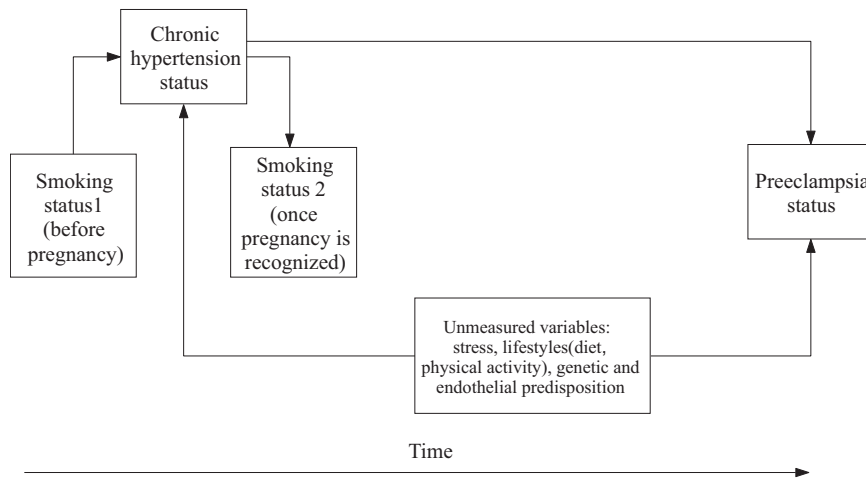


Fig. 4 Collider bias occurs when adjusting the analysis for a variable that is a common effect both of the exposure and of a cause of the outcome. Arrows indicate the direction of causality. Gray line represents the association of interest.

controlled. All these scenarios highlight the importance of following the entire cohort of women throughout their pregnancy, rather than only those who give birth, and to account for the reasons and characteristics of women who exit pregnancy at various times.

The loss of women at-risk affects the likelihood of an individual being sampled, and therefore affects selection. This is the case for studies where the population is selected on birth register information or retrospective studies which use prevalent cases [37]. Losses before, and after 20 weeks of GA need to be carefully considered in cohort studies as most traditional survival approaches require non-informative censoring, that is, censored women have the same probability of experiencing PE than those remaining in the study, which is not the case here.

Likewise, GH is defined as high BP after week 20, in the absence of proteinuria, organ and uteroplacental dysfunction. Despite it can be considered a milder form of PE, GH is usually considered as a different condition [1]. The incidence of GH versus PE increases earlier during pregnancy. As matter of fact, 50% of GH cases occur before the 35th week of GA [3]. Once GH is diagnosed, closer follow up visits, pharmacological treatment and, in some cases, labor induction before week 37 are offered to women. Hence, the prompt identification and treatment of GH early in pregnancy might impede the progression from GH to PE. Therefore, GH could be either a less severe form of PE or a competing event in the relation between smoking and PE.

Confounders

Temporal relations between variables and the period of observation are essential in identifying the role of a variable within a causal framework, to then determine whether to apply confounding control strategies [32]. There are controversial findings regarding smoking being a risk factor for CH [54–56]. If smoking increases the risk of masked hypertension, [57], and the day-time BP by up to 5–6 mmHg when smoking two cigarettes per hour [58], it could increase the likelihood of incident hypertension [56] and the chance that a patient can be diagnosed with hypertension. Once CH is present, smokers are advised to quit, especially those planning to get pregnant.

In Fig. 4, CH can work as a mediator in the relationship between smoking at time 1 and PE, this would be the case of a woman that starts smoking in adolescence, after which is diagnosed with CH and after that becomes pregnant and develop PE. If CH is considered a mediator, one can assess the direct effect of smoking on the risk of PE, i.e., the effect that is not mediated through the effect of smoking on CH. The direct effect of smoking would be

correctly estimated if: (1) there is no interaction between smoking and CH; that is, the effect of smoking on PE risk is the same in both CH + and CH- patients, and (2) the common causes (confounders) of the relation between the mediator (CH) and the outcome (PE) are controlled [59]. If common causes of both CH and PE are not controlled, the association between the mediator and the outcome is confounded and the mediator becomes a collider. Adjusting for a collider could create a biased association, as previously explained. CH is typically treated as a confounder using exclusion or another adjustment strategy [15, 60]. CH acts as a confounder in the relationship between smoking at time 2 and PE. That is, having a CH diagnosis subsequently affects both smoking status and the risk of PE. In such scenarios, the attempt to control for CH is reasonable.

All these temporal relations are usually unaccounted for in live-birth observational studies because the time of onset for each condition is not registered, and smoking is usually considered a time-fixed variable. Indeed, dependencies and temporal relationships between variables in a DAG are assumed to be the same for all women in the study population. Estimating direct effect and modeling all potential pathways is cumbersome due to data availability restrictions. In addition, birth registers do not usually include information about potential unmeasured confounders like lifestyle factors, which could also contribute to the counter-intuitive association.

Misclassifications

A source of misclassification bias is the inclusion of proteinuria for the diagnosis of PE, which is neither 100% sensitive nor specific. In addition, a young woman with previous unrecognized CH, might become normotensive beginning pregnancy, since BP usually drops during the first trimesters. Under current classification of hypertensive disorders in pregnancy, high BP values before 20th week of GA indicate the presence of CH, after which a new onset proteinuria or maternal organ dysfunction consistent with PE leads to diagnosis of PE superimposed on CH [1]. Persistent high BP beyond week 12 postpartum, in a woman with a previous diagnosed GH, indicates an unrecognized CH. Thus, any type of hypertension in pregnancy recorded at birth can be reclassified during postpartum period. Regarding the exposure, smoking status as a dichotomous variable, even if recorded at the beginning of pregnancy, ignores any cumulative effects from previous tobacco's exposure. Information bias can be induced if the exposure is self-reported, if those who quit are treated as never exposed, or if those smoking even one cigarette in life are treated as smoker.

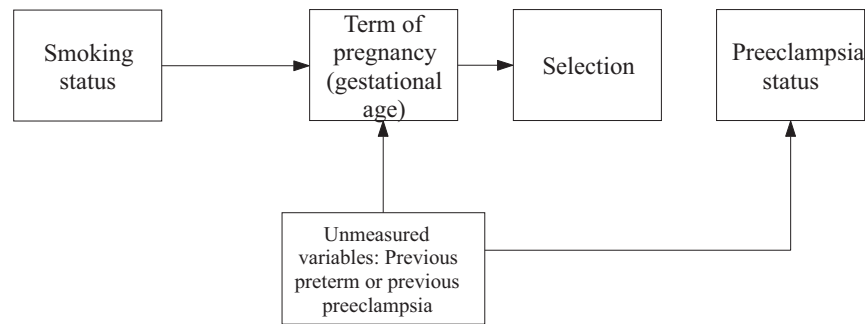


Fig. 5 Collider bias occurs when stratifying or adjusting the analysis for a variable that is a common effect both of the exposure and of a cause of the outcome. Arrows indicate the direction of causality. Gray line represents the association of interest.

The role of gestational age

GA has been used in stratified analysis. Adjustment for GA as a confounder or as a mediating variable will lead to bias when analyzing the association between prenatal factors and neonatal outcomes [61]. Under the smoking-PE causal hypothesis, GA is unlikely to create confounding as it is not a cause of smoking, unless smoking cessation increases with GA. However, the act of quitting smoking can only be present among smokers. One may argue that a certain GA is required to reach a diagnosis of PE according to current medical definitions, so that GA would be a necessary cause or a mediator. Luque-Fernandez analyzed GA as a mediator and reported an OR > 1 [28, 62]. Cases of postpartum PE diagnosed within 48 h and up to 6 weeks after birth were not included in that study. GA is the time frame from “conception”, or usually from the last menstrual period to a certain outcome (e.g., PE, delivery). In general, time window-at-risk to develop PE is equal to the duration of gestation. Therefore, GA can be treated as a time variable in survival analysis.

Dichotomized GA is more likely to be a collider, because it can be a common effect of smoking, and of an unmeasured cause of PE [63, 64] as shown in Fig. 5. Stratification on GA is implicit when clinical researchers define PE as preterm (<37 weeks) or term (≥37 weeks) [15], or as early-onset (<34 weeks) or late-onset (≥34 weeks) [65]. Preterm deliveries might preclude the incidence of PE at term. Early delivery due to other cause is then a competing event for most PE cases occurring at term. Maternal smoking and unmeasured confounders can be both associated with preterm delivery and PE [28]. This issue also applies when matching cases and controls by preterm/term delivery. Case-control designs are vulnerable to collider stratification bias, as selection into the study population is determined by the PE status. If smoking also influences the selection or the cause of the selection (e.g., GA), bias appears or is exacerbated [26].

UNPACKING THE PUZZLING ASSOCIATION

Similar “paradoxes” can appear in any medical field when the control for a confounder, mediator or a collider ignores the direction and temporal relations between variables. Some strategies to extricate the role of selection-bias in the presence of paradoxical effects are outlined below; however, further studies are warranted to evaluate the advantage of concurrent usage of these strategies.

1. To draw the causal diagram of the study question during the design, analysis and when interpreting the study results. The use of graphical tools such as DAGs [26] can help to identify sources of bias [66]. The temporal relationship between variables should be accounted for, as the role of a variable could change over time. In studies where the temporal relation between variables is not correctly accounted for,

bias due to reverse causation must be acknowledged. Joint models including survival and longitudinal data simultaneously are now available in the field, however they do not allow for the application of different DAGs on different subsets of the study population and then modelling all pathways simultaneously.

2. Cohort studies starting before conception or very early in pregnancy and the use of the fetus at risk approach are preferred over studies based on birth registers. In general, paradoxical effects from observational studies, starting after the onset of exposure, should be interpreted with care [67–69].
3. To specify whether total, indirect, or direct causal effects are relevant to the research question. When assessing the direct or indirect effect of an exposure, a mediator may become a collider if there is unmeasured confounder between the mediator and the outcome. It is preferable to exclude the collider/mediator in the analysis if total effect is of interest.
4. To eliminate confounding, adjusting for a common cause is generally needed, while adjusting for a common effect is not. Inverse probability weighting and g-methods may be advantageous for addressing covariates without generating collider bias [70].
5. Dealing with GA is cumbersome in the analysis [61]. The use of the fetuses-at-risk approach [71], including GA as a time variable, and using a competing risk model [72, 73] are strategies to be considered for outcomes that occur before birth [74]. There are valid analytical approaches to manage mediators [75–77]. However, there is no consensus on the correct strategy to handle GA [78–80]. Sensitivity analysis to estimate the magnitude of the bias [81], and strategies to include mediators and competing events in survival analysis [82] might be necessary.

CONCLUSION

“One size does not fit all” seems to apply when trying to explain the association between smoking and PE. Several sources of bias could explain this association including but not limited to eligibility criteria, early losses, competing events, the definition and misclassification of preeclampsia, inadequate adjustment, measurement errors of smoking, and unmeasured confounders. In addition, causal explanations based on studies where the temporal relation between variables is not guaranteed for each participant are challenging, since the same DAG’s structure and dynamics might not apply for the overall study population. More generally, estimating the average risk on a probabilistic approach for the determination of specific individual causal effects is frequently cumbersome [83]. To conclude, the reported biological protective effect of smoking on PE might be controverted by epidemiological reasoning.

DATA AVAILABILITY

Data are contained within the paper.

REFERENCES

- Poon LC, Shennan A, Hyett JA, Kapur A, Hadar E, Divakar H, et al. The International Federation of Gynecology and Obstetrics (FIGO) initiative on pre-eclampsia: A pragmatic guide for first-trimester screening and prevention. *Int J Gynaecol Obstet.* 2019;145:1–33.
- Magee LA, Nicolaidis KH, von Dadelszen P. Preeclampsia. *N Engl J Med.* 2022;386:1817–32.
- Olié V, Moutengou E, Grave C, Deneux-Tharoux C, Regnault N, Kretz S, et al. Prevalence of hypertensive disorders during pregnancy in France (2010–2018): The Nationwide CONCEPTION Study. *J Clin Hypertens (Greenwich).* 2021;23:1344–53.
- Burton GJ, Redman CW, Roberts JM, Moffett A. Pre-eclampsia: pathophysiology and clinical implications. *BMJ* 2019;366:12381.
- Perry H, Khalil A, Thilaganathan B. Preeclampsia and the cardiovascular system: An update. *Trends Cardiovasc Med.* 2018;28:505–13.
- Melchiorre K, Giorgione V, Thilaganathan B. The placenta and preeclampsia: villain or victim? *Am J Obstet Gynecol.* 2022;226:S954–S62.
- Zhang J, Klebanoff MA, Levine RJ, Puri M, Moyer P. The puzzling association between smoking and hypertension during pregnancy. *Am J Obstet Gynecol.* 1999;181:1407–13.
- Engel SM, Scher E, Wallenstein S, Savitz DA, Alsaker ER, Trostad L, et al. Maternal active and passive smoking and hypertensive disorders of pregnancy: risk with trimester-specific exposures. *Epidemiology* 2013;24:379–86.
- Conde-Agudelo A, Belizán JM. Risk factors for pre-eclampsia in a large cohort of Latin American and Caribbean women. *BJOG* 2000;107:75–83.
- Marcoux S, Brisson J, Fabia J. The effect of cigarette smoking on the risk of preeclampsia and gestational hypertension. *Am J Epidemiol.* 1989;130:950–7.
- Wei J, Liu CX, Gong TT, Wu QJ, Wu L. Cigarette smoking during pregnancy and preeclampsia risk: a systematic review and meta-analysis of prospective studies. *Oncotarget* 2015;6:43667–78.
- England L, Zhang J. Smoking and risk of preeclampsia: a systematic review. *Front Biosci.* 2007;12:2471–83.
- Wang J, Yang W, Xiao W, Cao S. The association between smoking during pregnancy and hypertensive disorders of pregnancy: A systematic review and meta-analysis. *Int J Gynaecol Obstet.* 2022;157:31–41.
- England LJ, Levine RJ, Qian C, Morris CD, Sibai BM, Catalano PM, et al. Smoking before pregnancy and risk of gestational hypertension and preeclampsia. *Am J Obstet Gynecol.* 2002;186:1035–40.
- Wikström AK, Stephansson O, Cnattingius S. Tobacco use during pregnancy and preeclampsia risk: effects of cigarette smoking and snuff. *Hypertension* 2010;55:1254–9.
- Conde-Agudelo A, Althabe F, Belizán JM, Kafury-Goeta AC. Cigarette smoking during pregnancy and risk of preeclampsia: a systematic review. *Am J Obstet Gynecol.* 1999;181:1026–35.
- Lewandowska M, Więckowska B. The Influence of Various Smoking Categories on The Risk of Gestational Hypertension and Pre-Eclampsia. *J Clin Med.* 2020;9:1743.
- Ekblad MO, Gissler M, Korhonen PE. New theory about the pathophysiology of preeclampsia derived from the paradox of positive effects of maternal smoking. *J Hypertens.* 2022;40:1223–30.
- Bakker R, Steegers EA, Mackenbach JP, Hofman A, Jaddoe VW. Maternal smoking and blood pressure in different trimesters of pregnancy: the Generation R study. *J Hypertens.* 2010;28:2210–8.
- Ambrose JA, Barua RS. The pathophysiology of cigarette smoking and cardiovascular disease: an update. *J Am Coll Cardiol.* 2004;43:1731–7.
- Virdis A, Giannarelli C, Neves MF, Taddei S, Ghiadoni L. Cigarette smoking and hypertension. *Curr Pharm Des.* 2010;16:2518–25.
- Karumanchi SA, Levine RJ. How does smoking reduce the risk of preeclampsia? *Hypertension* 2010;55:1100–1.
- Laule CF, Wing CR, Odean EJ, Wilcox JA, Gilbert JS, Regal JF. Effect of nicotine on placental ischemia-induced complement activation and hypertension in the rat. *J Immunotoxicol.* 2017;14:235–40.
- Hernández-Díaz S, Schisterman EF, Hernán MA. The birth weight “paradox” uncovered? *Am J Epidemiol.* 2006;164:1115–20.
- Wilcox AJ. On the importance—and the unimportance—of birthweight. *Int J Epidemiol.* 2001;30:1233–41.
- Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15:615–25.
- Lisonkova S, Joseph KS. Left truncation bias as a potential explanation for the protective effect of smoking on preeclampsia. *Epidemiology* 2015;26:436–40.
- Luque-Fernandez MA, Zoega H, Valdimarsdottir U, Williams MA. Deconstructing the smoking-preeclampsia paradox through a counterfactual framework. *Eur J Epidemiol.* 2016;31:613–23.
- Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology* 2009;20:488–95.
- Rothman KJ, Greenland S, Lash TL (eds). *Modern Epidemiology*. 3rd ed. Philadelphia: Wolters Kluwer and Lippincott Williams & Wilkins; 2008.
- Pearl J. *Causal Diagrams and the Identification of Causal Effects*. In: Pearl J (eds). *Causality. Models, Reasoning, and Inference*. 2nd ed. New York: Cambridge University Press; 2009. pp 65–102.
- Hernán MA, Hernández-Díaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am J Epidemiol.* 2002;155:176–84.
- Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, et al. Illustrating bias due to conditioning on a collider. *Int J Epidemiol.* 2010;39:417–20.
- Munafò MR, Tilling K, Taylor AE, Evans DM, Davey, Smith G. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol.* 2018;47:226–35.
- Luque-Fernandez MA, Schomaker M, Redondo-Sanchez D, Jose Sanchez Perez M, Vaidya A, Schnitzer ME. Educational Note: Paradoxical collider effect in the analysis of non-communicable disease epidemiological data: a reproducible illustration and web application. *Int J Epidemiol.* 2019;48:640–53.
- Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10:37–48.
- Chiu YH, Stensrud MJ, Dahabreh IJ, Rinaldo P, Diamond MP, Hsu J, et al. The Effect of Prenatal Treatments on Offspring Events in the Presence of Competing Events: An Application to a Randomized Trial of Fertility Therapies. *Epidemiology* 2020;31:636–43.
- Tanaka M, Tsujimoto Y, Goto K, Kumahara K, Onishi S, Iwanari S, et al. Preeclampsia before 20 weeks of gestation: a case report and review of the literature. *CEN Case Rep.* 2015;4:55–60.
- Seely EW, Ecker J. Chronic hypertension in pregnancy. *Circulation* 2014;129:1254–61.
- Yang Q, Wen SW, Smith GN, Chen Y, Krewski D, Chen XK, et al. Maternal cigarette smoking and the risk of pregnancy-induced hypertension and eclampsia. *Int J Epidemiol.* 2006;35:288–93.
- Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol.* 2008;167:492–9.
- Quenby S, Gallos ID, Dhillon-Smith RK, Podesek M, Stephenson MD, Fisher J, et al. Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss. *Lancet* 2021;397:1658–67.
- Winter E, Wang J, Davies MJ, Norman R. Early pregnancy loss following assisted reproductive technology treatment. *Hum Reprod.* 2002;17:3220–3.
- Nielsen A, Hannibal CG, Lindekilde BE, Tolstrup J, Frederiksen K, Munk C, et al. Maternal smoking predicts the risk of spontaneous abortion. *Acta Obstet Gynecol Scand.* 2006;85:1057–65.
- Bulletins—Gynecology ACoOGCoP. ACOG Practice Bulletin No. 200: Early Pregnancy Loss. *Obstet Gynecol.* 2018;132:e197–e207.
- Everett C. Incidence and outcome of bleeding before the 20th week of pregnancy: prospective study from general practice. *BMJ* 1997;315:32–4.
- 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;319:189–94.
- Modvig J, Schmidt L, Damsgaard MT. Measurement of total risk of spontaneous abortion: the virtue of conditional risk estimation. *Am J Epidemiol.* 1990;132:1021–38.
- Stovitz SD, Banack HR, Kaufman JS. ‘Depletion of the susceptibles’ taught through a story, a table and basic arithmetic. *BMJ Evid Based Med.* 2018;23:199.
- Kinlaw AC, Buckley JP, Engel SM, Poole C, Brookhart MA, Keil AP. Left Truncation Bias to Explain the Protective Effect of Smoking on Preeclampsia: Potential, But How Plausible? *Epidemiology* 2017;28:428–34.
- Windham GC, Swan SH, Fenster L. Parental cigarette smoking and the risk of spontaneous abortion. *Am J Epidemiol.* 1992;135:1394–403.
- Pipkin FB, Consortium GoP. Smoking in moderate/severe preeclampsia worsens pregnancy outcome, but smoking cessation limits the damage. *Hypertension* 2008;51:1042–6.
- Ananth CV, Vintzileos AM. Maternal-fetal conditions necessitating a medical intervention resulting in preterm birth. *Am J Obstet Gynecol.* 2006;195:1557–63.
- Gao K, Shi X, Wang W. The life-course impact of smoking on hypertension, myocardial infarction and respiratory diseases. *Sci Rep.* 2017;7:4330.
- Sohn K. Relationship of Smoking to Hypertension in a Developing Country. *Glob Heart* 2018;13:285–92.
- Kaplan RC, Baldoni PL, Strizich GM, Pérez-Stable EJ, Saccone NL, Peralta CA, et al. Current Smoking Raises Risk of Incident Hypertension: Hispanic Community Health Study-Study of Latinos. *Am J Hypertens.* 2021;34:190–7.
- Omboni S. Smoking and hypertension: what is behind the mask? *J Hypertens.* 2020;38:1029–30.

58. Gropelli A, Giorgi DM, Omboni S, Parati G, Mancia G. Persistent blood pressure increase induced by heavy smoking. *J Hypertens*. 1992;10:495–9.
59. Kaufman JS, Macleahose RF, Kaufman S. A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation. *Epidemiol Perspect Innov*. 2004;1:4.
60. Flegal KM, Graubard BI, Yi SW. Comparative effects of the restriction method in two large observational studies of body mass index and mortality among adults. *Eur J Clin Invest*. 2017;47:415–21.
61. Wilcox AJ, Weinberg CR, Basso O. On the pitfalls of adjusting for gestational age at birth. *Am J Epidemiol*. 2011;174:1062–8.
62. Ananth CV, Schisterman EF. Confounding, causality, and confusion: the role of intermediate variables in interpreting observational studies in obstetrics. *Am J Obstet Gynecol*. 2017;217:167–75.
63. Liu B, Xu G, Sun Y, Qiu X, Ryckman KK, Yu Y, et al. Maternal cigarette smoking before and during pregnancy and the risk of preterm birth: A dose-response analysis of 25 million mother-infant pairs. *PLoS Med*. 2020;17:e1003158.
64. Moore E, Blatt K, Chen A, Van Hook J, DeFranco EA. Relationship of trimester-specific smoking patterns and risk of preterm birth. *Am J Obstet Gynecol*. 2016;215:109.e1–6.
65. Raymond D, Peterson E. A critical review of early-onset and late-onset pre-eclampsia. *Obstet Gynecol Surv*. 2011;66:497–506.
66. Banack HR, Kaufman JS. The “obesity paradox” explained. *Epidemiology* 2013;24:461–2.
67. Lajous M, Banack HR, Kaufman JS, Hernán MA. Should patients with chronic disease be told to gain weight? The obesity paradox and selection bias. *Am J Med*. 2015;128:334–6.
68. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016;79:70–5.
69. Neophytou AM, Kioumourtzoglou MA, Goin DE, Darwin KC, Casey JA. Educational note: addressing special cases of bias that frequently occur in perinatal epidemiology. *Int J Epidemiol*. 2021;50:337–45.
70. Samuel M, Batomen B, Rouette J, Kim J, Platt RW, Brophy JM, et al. Evaluation of propensity score used in cardiovascular research: a cross-sectional survey and guidance document. *BMJ Open*. 2020;10:e036961.
71. 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;43:1154–62.
72. Dignam JJ, Zhang Q, Kocherginsky M. The use and interpretation of competing risks regression models. *Clin Cancer Res*. 2012;18:2301–8.
73. Haller B, Schmidt G, Ulm K. Applying competing risks regression models: an overview. *Lifetime Data Anal*. 2013;19:33–58.
74. Harmon QE, Basso O, Weinberg CR, Wilcox AJ. Two denominators for one numerator: the example of neonatal mortality. *Eur J Epidemiol*. 2018;33:523–30.
75. VanderWeele TJ, Mumford SL, Schisterman EF. Conditioning on intermediates in perinatal epidemiology. *Epidemiology* 2012;23:1–9.
76. Auger N, Naimi AI, Fraser WD, Healy-Profitós J, Luo ZC, Nuyt AM, et al. Three alternative methods to resolve paradoxical associations of exposures before term. *Eur J Epidemiol*. 2016;31:1011–9.
77. Ananth CV, Brandt JS. A principled approach to mediation analysis in perinatal epidemiology. *Am J Obstet Gynecol*. 2022;226:24–32.e6.
78. Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology* 2003;14:300–6.
79. Hafeman DM. Confounding of indirect effects: a sensitivity analysis exploring the range of bias due to a cause common to both the mediator and the outcome. *Am J Epidemiol*. 2011;174:710–7.
80. Smith LH, VanderWeele TJ. Bounding Bias Due to Selection. *Epidemiology* 2019;30:509–16.
81. VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology* 2010;21:540–51.
82. Lapointe-Shaw L, Bouck Z, Howell NA, Lange T, Orchanian-Cheff A, Austin PC, et al. Mediation analysis with a time-to-event outcome: a review of use and reporting in healthcare research. *BMC Med Res Methodol*. 2018;18:118.
83. Merlo J. Invited commentary: multilevel analysis of individual heterogeneity—a fundamental critique of the current probabilistic risk factor epidemiology. *Am J Epidemiol*. 2014;180:208–12.

AUTHOR CONTRIBUTIONS

Design and conception: MR-L. Conducted the review of literature: MR-L, MFE, JM, JK. Figure design: MR-L, JK. Wrote the first draft of the paper: MR-L. Reviewed all versions of the paper and contributed to the structure of the review: MR-L, MFE, JM, JK. Editing of the paper, and have approved the final draft: MR-L, MFE, JM, JK.

FUNDING

This work was partially supported by the Swedish Research Council, Grant/Award Number: 2017-01321 and 2013-2484 (JM). Open access funding provided by Lund University.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Merida Rodriguez-Lopez.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023