International Journal of Epidemiology, 2015, 345–354
doi: 10.1093/ije/dyu249
Advance Access Publication Date: 19 January 2015
Original article

Education Corner

Bias from conditioning on live birth in pregnancy cohorts: an illustration based on neurodevelopment in children after prenatal exposure to organic pollutants

Zeyan Liew,[1]* Jørn Olsen,[1,2] Xin Cui,[1] Beate Ritz[1,3] and Onyebuchi A Arah[1,4,5]

[1]Department of Epidemiology, Fielding School of Public Health, University of California Los Angeles (UCLA), Los Angeles, CA, USA, [2]Department of Public Health, Section for Epidemiology, University of Aarhus, Aarhus, Denmark, [3]Department of Neurology, Geffen School of Medicine, UCLA, Los Angeles, CA, USA, [4]Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands and [5]UCLA Center for Health Policy Research, Los Angeles, CA, USA

*Corresponding author. Department of Epidemiology, UCLA Fielding School of Public Health, 650 Charles E. Young Drive, Los Angeles, CA 90095-1772, USA. E-mail: zeyanliew@ucla.edu

Accepted 21 August 2014

Abstract

Only 60–70% of fertilized eggs may result in a live birth, and very early fetal loss mainly goes unnoticed. Outcomes that can only be ascertained in live-born children will be missing for those who do not survive till birth. In this article, we illustrate a common bias structure (leading to 'live-birth bias') that arises from studying the effects of prenatal exposure to environmental factors on long-term health outcomes among live births only in pregnancy cohorts. To illustrate this we used prenatal exposure to perfluoroalkyl substances (PFAS) and attention-deficit/hyperactivity disorder (ADHD) in school-aged children as an example. PFAS are persistent organic pollutants that may impact human fecundity and be toxic for neurodevelopment. We simulated several hypothetical scenarios based on characteristics from the Danish National Birth Cohort and found that a weak inverse association may appear even if PFAS do not cause ADHD but have a considerable effect on fetal survival. The magnitude of the negative bias was generally small, and adjusting for common causes of the outcome and fetal loss can reduce the bias. Our example highlights the need to identify the determinants of pregnancy loss and the importance of quantifying bias arising from conditioning on live birth in observational studies.

Key words: Bias analysis, live-birth bias, reproductive epidemiology, birth cohort, prenatal exposure, perfluoroalkyl substances, attention-deficit/hyperactivity disorder

Key Messages

- 'Live-birth bias' may arise in studies using pregnancy cohorts to investigate the impact of prenatal exposures on health outcomes that manifest only after births.
- Simulation analyses based on information from the Danish National Birth Cohort and several hypothetical scenarios suggest that weak inverse associations between prenatal PFAS and ADHD in children can be observed even if PFAS do not cause or prevent ADHD but have considerable effects on fetal survival.
- The magnitude of bias from 'conditioning on fetal survival' is generally small; one way to reduce this bias is to adjust for measured risk factors of the study outcome that also have an impact on fetal survival.
- It is important to improve our understanding of the determinants of fetal loss and collect such data in pregnancy exposure studies.
- Quantitative bias analyses is important to estimate the size of the impact of a 'live-birth bias' in observational studies.

Introduction

In life-course and reproductive epidemiology, we often aim to study whether exposures induce fetal programming as a function of the time the population is at risk, and that time starts at conception. However, the total number of conceptions in the source population is generally unknown because of the complex selection phenomena in human reproduction such that only 60–70% of all fertilized eggs will likely result in a live birth.[1,2] Previous studies that relied on the rise of human chorionic gonadotropin (hCG) as a biomarker for detecting pregnancies early estimated the incidence of early pregnancy loss (EPL; defined as fetal loss after implantation and before pregnancy was detected clinically) to range around 20–30% in cohorts of pregnancy planners.[1,3] Using hCG as a biomarker however does not allow for pregnancy loss detection before implantation, and planned pregnancy might be different from unplanned pregnancy; thus the rates of EPL in the general population could be higher.[2] Genetic, hormonal and immunological factors,[4,5] as well as exposure to environmental chemicals such as endocrine-disrupting compounds, heavy metals and cigarette smoke,[6,7] affect infertility and pregnancy loss in humans.

Health effects from exposures in pregnancy that can only be ascertained in live-born children will be missed if exposures contribute to the abortion of fetuses, thus reducing the number of exposed life-born conceptions. This fact has become well known in studies of congenital malformations or birth defects which are thus considered in terms of prevalence measures,[8] since it is known that exposure that reduces fetal survival disproportionately among malformed fetuses will result in lower ('preventive') prevalence measures among the exposed. Basically, exposure moves some malformations from being identifiable into a window of invisibility. Few studies have employed quantitative methods to examine the impact on effect estimates for childhood or long-term health outcomes as a consequence of pregnancy exposures when exposure can cause fetal loss.

Several recent epidemiological studies observed an unexpected inverse association between prenatal levels of perfluoroalkyl substances (PFAS) and risk of attention-deficit/hyperactivity disorder (ADHD) in children.[9–12] PFAS are man-made synthetic chemicals that have been extensively used as surfactants in industrial and commercial applications such as food packaging material, non-stick pan coatings, and personal care products. PFAS are extremely persistent in the environment and in humans; for example, the estimated biological half-life is around 4 to 8 years for common PFAS. PFAS can cross the placental barrier, exposing the fetus, and animal studies showed that PFAS are neurotoxic during development.[13,14] However, two previous reports based on subsets of the Danish National Birth Cohort (DNBC) unexpectedly showed weak inverse associations between prenatal PFAS levels and behavioural problems[9] or ADHD diagnosis.[10] Specifically, children born to mothers in the highest quartile of perfluorooctanesulfonate (PFOS) levels, the most common PFAS, were at lower risk of receiving a diagnosis of ADHD [Risk Ratio = 0.79; 95% confidence interval (CI) 0.64-0.98]. A smaller study in Sweden estimated an odds ratio of 0.81 (95% CI 0.50–1.32) for PFOS in umbilical cord serum and ADHD in children.[11] Another longitudinal study, conducted in a community that for decades was highly exposed to perfluorooctanoate (PFOA) through contaminated drinking water, also reported higher levels of *in utero* exposure to PFOA to be associated with fewer, not more, ADHD symptoms in offspring.[12] There is, however, no biological explanation for PFAS protecting children or the developing brain from ADHD. Error due to chance is a possible explanation, but other potential bias should also be considered.

Prenatal exposures to PFAS have been suggested to increase the incidence of fetal resorptions and to cause fetal

and neonatal deaths in animals.[15–17] In humans, PFAS were shown to interfere with sex hormone and thyroid hormone homeostasis,[18,19] and higher prenatal levels of PFAS were associated with reduced fecundity[20,21] and increased risk for miscarriage.[22] In addition, PFAS may have an impact on sperm morphology, semen quality and reproductive hormone levels in men,[23] which could result in reduced fecundity and elevate the risk for early pregnancy loss.

In this article, we first use directed acyclic graphs to illustrate the common bias structure, namely conditioning on live-birth status when studying the impact of prenatal environmental exposures on long-term health outcomes. Next, we use PFAS and ADHD as an illustrative example to study the direction and magnitude of the bias; this bias structure could be generalized to many other environmental threats and outcomes that share the causal and bias structure described here. We simulated a few assumed scenarios based on information taken from the Danish National Birth Cohort, and investigate whether conditioning on live-born status may induce sufficiently large bias to explain the unexpected inverse associations between prenatal PFAS exposures and ADHD in our and previous studies.

Methods

Bias structure

We use directed acyclic graphs (DAGs) to present the structural relation between variables. The basic set of rules in utilizing DAGs has been described elsewhere.[24] Two major sources of biasing paths are uncontrolled confounding (Supplementary Figure 1a, available as Supplementary data at *IJE* online), i.e. the failure to control for a confounder such as a common cause, and conditioning on a collider or a common effect of the two variables under study, thus opening up an otherwise closed path (Supplementary Figure 1b, available as Supplementary data at *IJE* online).

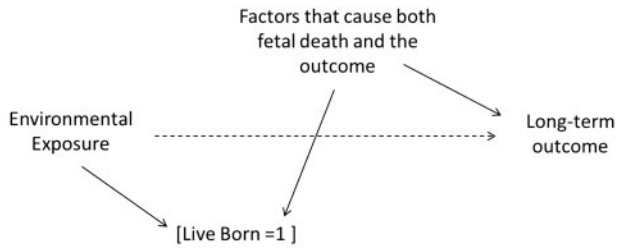

Figure 1. A common bias structure in studies of prenatal exposure to environmental factors on long-term health outcomes where the environmental exposure causes fetal losses. Only the fetuses that survived (indicated by live born equal to 1) can be ascertained for the long-term outcome. Conditioning on live-born status opens up a collider path from environmental exposure to outcome via other uncontrolled common causes of fetal survival and the outcome.

A well-known example of collider bias is selection bias resulting from study participants' differential (in terms of outcome and exposure) non-response or loss to follow-up.[25] Competing risk and survivor bias have also been previously conceptualized as forms of collider bias.[26]

Figure 1 shows the common bias structure we propose for studies of prenatal exposure to environmental factors and health outcomes that manifest after birth when the environmental exposure also causes fetal loss. Since outcomes that can only be ascertained in live-born children will be missing for all non-surviving fetuses, and if exposures also contribute to loss, when we condition on live-birth status (as is typical in perinatal and paediatric cohort studies) we induce a 'collider(-stratification) bias' in the presence of uncontrolled common causes of fetal death and the health outcome after birth. This type of bias could be called 'live-birth bias' in perinatal and paediatric epidemiology. This live-birth bias is distinct from and can coexist with the birthweight paradox in smoking-low-birthweight-mortality studies,[27] and with the fetuses-at-risk bias in maternal-risk-factor-preterm-mortality studies.[28] The latter collider-bias scenarios arise from conditioning on postnatal variables such as low birthweight, not from missing data on unobserved fetal loss that could have been the common consequence of the prenatal exposure under study and some unknown common causes of the postnatal outcome under study and fetal loss.

Possible theories for PFAS and ADHD

We propose three scenarios based on what we consider realistic, given existing knowledge regarding associations between PFAS, ADHD and fetal survivals, and measured as well as unmeasured risk factors of ADHD. However, we emphasize that the proposed scenarios are hypothetical, i.e. they are based on the assumption that PFAS can influence human fecundity and increase risk of fetal death. While research indicated that PFAS may impact fetal loss,[15–17,20–22] more studies and data are still needed to evaluate whether PFAS lead to fetal death in humans. In the first scenario (Figure 2a), there is no causal relation between PFAS and ADHD. However, exposure to PFAS decreases the chance of conception (C) and fetal survival before clinical detection of pregnancy (S1) or a live birth (S2). Based on the literature, we introduce a variable R which represents a set of known risk factors for ADHD that also influence conception and fetal survival, specifically maternal socioeconomic status,[29] smoking[30] and psychological stress.[31] In addition, the variable U represents a set of unmeasured, possibly unknown, risk factors for ADHD that reduce fecundity and fetal survival. Some possible candidates for U include predisposition due to genetic

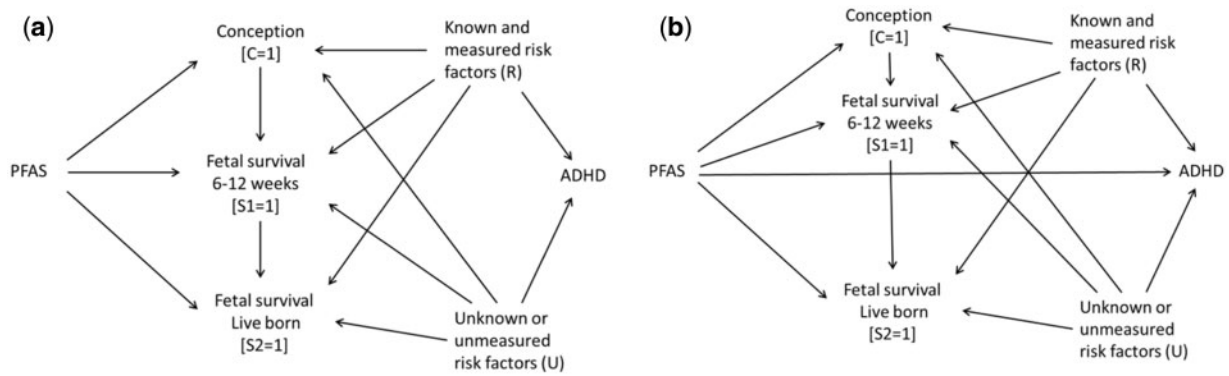

Figure 2. (a) DAG of scenario 1 where no causal relation between PFAS and ADHD in the simulated pregnancy cohort. However, conditioning on conception and fetal survival (C, S1 and S2) opens up biasing paths from PFAS to ADHD through risk factors of ADHD (R and U). (b) DAG of scenario 2 where a causal relation (direct effect) between PFAS and ADHD was assumed. Conditioning on C, S1 and S2 would induce similar collider biases to those described in (a).

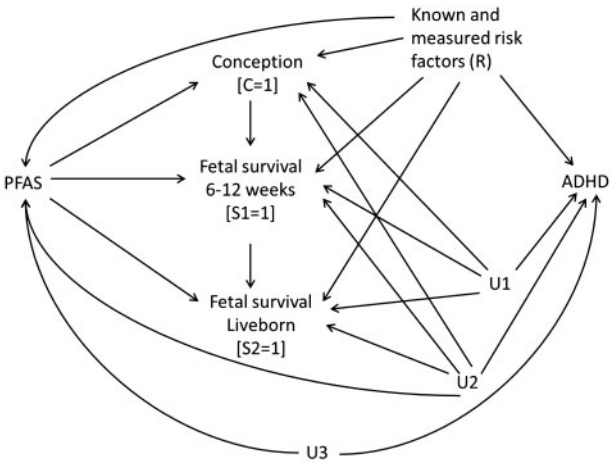

Figure 3. DAG of scenario 3 where no causal relation between PFAS and ADHD in the simulated pregnancy cohort. Conditioning on conception and fetal survival (C, S1 and S2) opens up biasing paths from PFAS to ADHD through risk factors of ADHD (R, U1 and U2). U2 and U3 are determinants of PFAS thus would additionally contribute confounding biases in this scenario.

factors or other environmental toxins.[32,33] In the second scenario (Figure 2b), the same set of variables was employed as in scenario 1 but here we now claim that prenatal exposure to PFAS causes a small increase in ADHD risk [odds ratio (OR) = 1.2 or 1.5] in children.

In addition, we also examine a third scenario (Figure 3) where we assumed no causal relation between PFAS and ADHD, and simulated three unknown or unmeasured risks factors (U1, U2 and U3). The third scenario is more complicated since it takes into consideration both collider bias and uncontrolled confounding. We simulated U1 to represent some unknown genetic predisposition that is a strong risk factor of ADHD (OR = 5) and has a strong effect on fetal survival (OR = 0.3). U2 represents unmeasured environmental and lifestyle factors that are moderately associated with ADHD (OR = 2) and fetal survival (OR = 0.5), and are also positively correlated with PFAS exposures (OR = 1.5). U3 represents potential unmeasured confounding factors that are common causes of PFAS and ADHD.

Simulations and statistical analysis

We used Monte Carlo techniques to perform simulations, basing our parameter priors on information from the Danish National Birth Cohort (DNBC). The DNBC is a nationwide cohort study that followed about 100 000 pregnancies and children with the aim to study pregnancy complications and diseases in offspring in relation to causes operating during pregnancy and early life (for details see Olsen *et al.*[34]). Pregnant women were recruited by general practitioners after their pregnancies were clinically recognized at around 6–12 weeks of gestation. According to maternal report, 75% of the enrolled pregnancies were planned and 25% were partially planned or not planned. ADHD diagnoses in children were ascertained based on ICD-10 diagnoses through linkage with admission records from all general and mental health hospitals in Denmark. Maternal PFAS levels during pregnancy in the DNBC have previously been ascertained and reported.[10,20] PFOS and PFOA are the most frequently detected PFAS, that is they were detected in all maternal plasma samples measured from the DNBC.[10]

We simulated a cohort consisting of 92 000 live-born singletons, i.e. the number in the DNBC with ~3% of all children developing ADHD. We assumed that about 20% of EPL remained undocumented and thus these pregnancies were not enrolled in the DNBC, and that 5% of fetal deaths occurred after cohort enrolment as documented previously in studies reporting on miscarriages and abortions in the DNBC.[35] We compared women with PFAS levels in the highest quartile with all others, thus generating a binary PFAS variable with the prevalence of exposure set as 25%. In all three scenarios, fixed priors were used for the prevalence of known risk factors (R) and the relation

between R and ADHD or fetal loss. In addition, a range of input levels were assigned to the unmeasured risks factors (U1, U2 and U3), including their presumed prevalence and the strength of their associations with ADHD and fetal loss. The list of variables with their assumed prevalence and strengths of associations are presented in Table 1.

We used SAS 9.3 (SAS Institute, Cary, NC, USA) to perform all simulations and analyses.[36] We generated binary values for the 'exogenous' variables (i.e. variables that have no arrow pointing towards them as shown in DAGs in Figure 2 and Figure 3) including PFAS in the first two scenarios, and known or unknown risk factors of ADHD (R and U's) in all scenarios, by random draws from independent Bernoulli distributions such that: PFAS $\sim$ B(1, 0.25), R $\sim$ B(1, 0.4) and U's $\sim$ B(1, P(U = 1)). For variables that have causal determinants (arrows that point towards them) such as conception (C), ADHD, and PFAS in the third scenario, simulations were conducted based on the following equations.

In scenarios 1 and 2:

$$C \sim B \quad (1, \quad (1/(1+\exp(-(\log(P(C=1)/(1-P(C=1)))+\log(OR_{C\text{-}PFAS})*PFAS+\log(OR_{C\text{-}R})*R+\log(OR_{C\text{-}U})*U)))))$$

In scenario 1:

$$ADHD \sim B \ (1, \ (1/(1+\exp(-(\log(P(ADHD=1)/(1-P(ADHD=1)))+\log(OR_{ADHD\text{-}R})*R+\log(OR_{ADHD\text{-}U})*U)))))$$

In scenario 2:

$$ADHD \sim B \ (1, \ (1/(1+\exp(-(\log(P(ADHD=1)/(1-P(ADHD=1)))+\log(OR_{ADHD\text{-}R})*R+\log(OR_{ADHD\text{-}U})*U+\log(OR_{ADHD\text{-}PFAS})*PFAS)))))$$

In scenario 3:

$$C \sim B \ (1, \ (1/(1+\exp(-(\log(P(C=1)/(1-P(C=1)))+\log(OR_{C\text{-}PFAS})*PFAS+\log(OR_{C\text{-}R})*R+\log(OR_{C\text{-}U1})*U1+\log(OR_{C\text{-}U2})*U2)))))$$
$$ADHD \sim B \ (1, \ (1/(1+\exp(-(\log(P(ADHD=1)/(1-P(ADHD=1)))+\log(OR_{ADHD\text{-}R})*R+\log(OR_{ADHD\text{-}U1})*U1+\log(OR_{ADHD\text{-}U2})*U2+\log(OR_{ADHD\text{-}U3})*U3)))))$$
$$PFAS \sim B \ (1, \ (1/(1+\exp(-(\log(P(PFAS=1)/(1-P(PFAS=1)))+\log(OR_{PFAS\text{-}R})*R+\log(OR_{PFAS\text{-}U2})*U2+\log(OR_{PFAS\text{-}U3})*U3)))))$$

For fetal survival in early pregnancy (S1) and live birth of infants (S2), we used conditional probabilities to assign values to S1 and S2 such as:

In scenarios 1 and 2:

$$S1 \sim B \quad (1, \quad (1/(1+\exp(-(\log(P(S1=1)/(1-P(S1=1)))+\log(OR_{S1\text{-}PFAS})*PFAS+\log(OR_{S1\text{-}R})*R+\log(OR_{S1\text{-}U})*U))))) \mid C=1$$

Table 1. Priors for the simulation studies

Variables[a]	Abbreviation	Prevalence	Specified relation (ORs) in scenarios 1 and 2	Specified relation (ORs) in scenario 3
Perfluoroalkyl substances	PFAS	25% exposed	no causal determinants	$OR_{PFAS\text{-}R} = 1.5$; $OR_{PFAS\text{-}U2} = 1.5$; $OR_{PFAS\text{-}U3} = 1.5, 3, 0.5, 0.3$
Conception	C	95%	$OR_{C\text{-}R} = 0.5$; $OR_{C\text{-}PFAS} = 0.8, 0.5$; $OR_{C\text{-}U} = 0.8, 0.5, 0.3$	$OR_{C\text{-}R} = 0.5$; $OR_{C\text{-}PFAS} = 0.8, 0.5$; $OR_{C\text{-}U1} = 0.3$; $OR_{C\text{-}U2} = 0.5$
Fetal survival in early gestation	S1	80% survived among all conceptions	$OR_{S1\text{-}R} = 0.5$; $OR_{S1\text{-}PFAS} = 0.8, 0.5$; $OR_{S1\text{-}U} = 0.8, 0.5, 0.3$	$OR_{S1\text{-}R} = 0.5$; $OR_{S1\text{-}PFAS} = 0.8, 0.5$; $OR_{S1\text{-}U1} = 0.3$; $OR_{S1\text{-}U2} = 0.5$
Fetal survival at birth[b]	S2	95% survived among those survived in early gestation	$OR_{S2\text{-}R} = 0.5$; $OR_{S2\text{-}PFAS} = 0.8, 0.5$; $OR_{S2\text{-}U} = 0.8, 0.5, 0.3$	$OR_{S2\text{-}R} = 0.5$; $OR_{S2\text{-}PFAS} = 0.8, 0.5$; $OR_{S2\text{-}U1} = 0.3$; $OR_{S2\text{-}U2} = 0.5$
Known and measured risk factors that impact on fetal survival	R	40%	no causal determinants	no causal determinants
Unknown or unmeasured risk factors	U (scenarios 1 and 2); U1,U2,U3 (scenario 3)	U = 20% or 40%; U1 = 10%, U2 = 20%, U3 = 20%	no causal determinants	no causal determinants
Attention-deficit/hyperactivity disorder	ADHD	3% among live-born	$OR_{ADHD\text{-}R} = 8$; $OR_{ADHD\text{-}U} = 2,5,10$; $OR_{ADHD\text{-}PFAS} = 1.2, 1.5$ (scenario2 only)	$OR_{ADHD\text{-}R} = 8$; $OR_{ADHD\text{-}U1} = 5$; $OR_{ADHD\text{-}U2} = 2$; $OR_{ADHD\text{-}U3} = 2, 4$

[a]All variables were generated as binary responses (1 = yes, 0 = no).
[b]About hypothetical 92 000 live-born were generated.

Table 2. Simulation results[a] of PFAS and ADHD in a hypothetical live-born birth cohort (scenario 1 assuming a true null effect of PFAS on ADHD)

$Pr(U=1)$[c]	OR_{ADHD-U}	OR_{C-U}, OR_{S1-U}, OR_{S2-U} =					
		0.8		0.5		0.3	
		Crude OR	Adjusted OR[b]	Crude OR	Adjusted OR[b]	Crude OR	Adjusted OR[b]
where OR_{C-PFAS}, $OR_{S1-PFAS}$, $OR_{S2-PFAS}$ = 0.8							
0.2	2	0.98	1.00 (0.90-1.09)	0.97	0.99 (0.89-1.10)	0.97	0.99 (0.89-1.10)
0.2	5	0.97	0.99 (0.91-1.08)	0.96	0.98 (0.90-1.08)	0.95	0.97 (0.88-1.07)
0.2	10	0.97	0.99 (0.92-1.06)	0.96	0.97 (0.90-1.05)	0.94	0.95 (0.87-1.04)
0.4	2	0.97	0.99 (0.91-1.09)	0.97	0.99 (0.90-1.09)	0.96	0.98 (0.87-1.09)
0.4	5	0.97	0.99 (0.92-1.06)	0.96	0.98 (0.90-1.05)	0.94	0.96 (0.87-1.04)
0.4	10	0.97	0.99 (0.94-1.05)	0.95	0.97 (0.91-1.03)	0.93	0.94 (0.87-1.02)
where OR_{C-PFAS}, $OR_{S1-PFAS}$, $OR_{S2-PFAS}$ = 0.5							
0.2	2	0.92	0.99 (0.89-1.11)	0.90	0.98 (0.87-1.09)	0.90	0.97 (0.86-1.08)
0.2	5	0.91	0.98 (0.89-1.07)	0.87	0.94 (0.85-1.04)	0.85	0.91 (0.82-1.00)
0.2	10	0.90	0.97 (0.90-1.06)	0.85	0.91 (0.84-0.99)	0.81	0.86 (0.78-0.94)
0.4	2	0.92	0.99 (0.89-1.10)	0.89	0.97 (0.86-1.08)	0.87	0.94 (0.83-1.06)
0.4	5	0.91	0.97 (0.90-1.06)	0.85	0.92 (0.84-1.00)	0.81	0.86 (0.77-0.95)
0.4	10	0.90	0.97 (0.91-1.04)	0.84	0.90 (0.83-0.96)	0.77	0.81 (0.74-0.89)

[a]Assume fixed priors for $Pr(R=1)=0.4$, OR_{C-R}, OR_{S1-R}, $OR_{S2-R}=0.5$, $OR_{ADHD-R}=8$.
[b]Adjusted for R (known and measured risk factors that impact on fetal survival).
[c] Prevalence of unknown or unmeasured risk factors.

$$S2 \sim B\ (1,\ (1/(1+\exp(-(\log(P(S2=1)/(1-P(S2=1)))+\log(OR_{S2\text{-}PFAS})*PFAS+\log(OR_{S2\text{-}R})*R+\log(OR_{S2\text{-}U})*U))))) \mid S1=1$$

In scenario 3:

$$S1 \sim B\ (1,\ (1/(1+\exp(-(\log(P(S1=1)/(1-P(S1=1)))+\log(OR_{PFAS\text{-}S1})*PFAS+\log(OR_{S1\text{-}R})*R+\log(OR_{S1\text{-}U1})*U1+\log(OR_{S1\text{-}U2})*U2))))) \mid C=1$$
$$S2 \sim B\ (1,\ (1/(1+\exp(-(\log(P(S2=1)/(1-P(S2=1)))+\log(OR_{S2\text{-}PFAS})*PFAS+\log(OR_{S2\text{-}R})*R+\log(OR_{S1\text{-}U1})*U1+\log(OR_{S1\text{-}U2})*U2))))) \mid S1=1$$

For each simulated dataset, we performed logistic regression analysis of ADHD status in children with prenatal PFAS as exposure, restricted to live births only ($S2=1$), further assuming that all children survived after birth until time of diagnosis and that there was no loss to follow-up preventing us from knowing the outcome status. The analysis was repeated for different simulation parameter values in each scenario. We reported odds ratios (ORs) and 95% simulation intervals using the 2.5, 50 and 97.5 percentile following 1000 simulation draws. We compared estimates with or without adjustment for known common causes of ADHD and fetal death. We calculated 'bias ratio' defined as assumed OR_{true} divided by $OR_{estimate}$ (using the point estimate) to examine magnitude of bias. When no bias is present then the bias ratio is equal to one, and greater departures from 1 indicate larger magnitude of bias.

We also conducted sensitivity analyses by varying the sample size of the simulated cohort (using 50% or 10% of the sample size of the DNBC), increasing the percentage of EPL, or assuming that PFAS only had an impact on EPL but not on conception or clinically recognized pregnancy loss.

Results

In scenario 1 where we assumed a null association among PFAS and ADHD to be true, we observed a 'protective' effect of PFAS on ADHD among live births after conditioning on fetal survival (Table 2). As expected, the magnitude of the inverse associations became stronger with increasingly stronger associations between PFAS and conceptions, early and late pregnancy loss. Moreover, protective associations also strengthened when the effect sizes for the influence of unmeasured risk factors on ADHD and pregnancy loss increased. The largest protective effect estimates were $OR_{crude}=0.77$ (95% simulation interval 0.70-0.83; bias ratio = 1.30) and $OR_{adjusted}=0.81$ (95% simulation interval 0.74-0.89; bias ratio = 1.23). Adjusted ORs were closer to the null than crude ORs, illustrating that controlling for known risk factors for ADHD and fetal deaths attenuates the negative bias slightly and moves the estimate closer to the true value. Results were similar when the prevalence of the unknown factors was assumed to be 0.2 or 0.4.

Table 3. Simulation results[a] of PFAS and ADHD in a hypothetical live-born birth cohort (scenario 2 assuming a true causal OR = 1.2 of PFAS on ADHD)

$Pr(U=1)$[c]	OR_{ADHD-U}	OR_{C-U}, OR_{S1-U}, OR_{S2-U} =					
		0.8		0.5		0.3	
		Crude OR	Adjusted OR[b]	Crude OR	Adjusted OR[b]	Crude OR	Adjusted OR[b]
where OR_{C-PFAS}, $OR_{S1-PFAS}$, $OR_{S2-PFAS}$ = 0.8							
0.2	2	1.17	1.20 (1.08-1.31)	1.16	1.19 (1.08-1.32)	1.16	1.19 (1.08-1.30)
0.2	5	1.15	1.18 (1.09-1.28)	1.14	1.17 (1.07-1.28)	1.13	1.16 (1.05-1.26)
0.2	10	1.14	1.17 (1.09-1.25)	1.12	1.15 (1.07-1.24)	1.10	1.13 (1.04-1.23)
0.4	2	1.16	1.19 (1.10-1.30)	1.16	1.19 (1.08-1.30)	1.15	1.18 (1.06-1.30)
0.4	5	1.15	1.18 (1.11-1.26)	1.13	1.16 (1.08-1.25)	1.11	1.14 (1.04-1.23)
0.4	10	1.13	1.17 (1.11-1.23)	1.11	1.14 (1.07-1.22)	1.08	1.11 (1.03-1.19)
where OR_{C-PFAS}, $OR_{S1-PFAS}$, $OR_{S2-PFAS}$ = 0.5							
0.2	2	1.10	1.19 (1.07-1.32)	1.08	1.17 (1.08-1.30)	1.08	1.16 (1.04-1.29)
0.2	5	1.08	1.17 (1.07-1.27)	1.04	1.12 (1.02-1.23)	1.01	1.08 (0.98-1.19)
0.2	10	1.06	1.14 (1.06-1.23)	1.00	1.07 (0.99-1.17)	0.95	1.02 (0.94-1.11)
0.4	2	1.09	1.19 (1.08-1.31)	1.06	1.16 (1.03-1.29)	1.05	1.13 (1.01-1.26)
0.4	5	1.07	1.16 (1.08-1.26)	1.01	1.09 (1.01-1.19)	0.96	1.03 (0.93-1.13)
0.4	10	1.05	1.14 (1.07-1.22)	0.98	1.05 (0.98-1.13)	0.90	0.96 (0.88-1.03)

[a]Assume fixed priors for $Pr(R=1)=0.4$, OR_{C-R}, OR_{S1-R}, $OR_{S2-R}=0.5$, $OR_{ADHD-R}=8$.
[b]Adjusted for R (known and measured risk factors that impacted on fetal survival).
[c] Prevalence of unknown or unmeasured risk factors.

In scenario 2, we assumed a moderate causal relation (true OR = 1.2 or 1.5) between PFAS and ADHD. Assuming the true OR of prenatal PFAS exposure and ADHD to be 1.2, conditioning on fetal survival resulted in a bias towards the null, but only few ORs crossed the null and fell below 1 (bias ratios range from 1.03 to 1.33 for crude ORs, and 1.00 to 1.25 for adjusted ORs) (Table 3). Similarly attenuated but still positive associations were seen when we assumed a stronger size effect of PFAS on ADHD (true OR = 1.5), but none of the point estimates fell below 1 (Table 4). Again, adjusting for known common causes of ADHD and fetal death removed some of the negative bias and moved effect estimates closer to the assumed true OR in this simulation study (bias ratios 1.03-1.38 for crude OR and 1.00-1.28 for adjusted OR).

Table 5 shows the results in scenario 3 (assumed true OR = 1 for PFAS and ADHD) in which we assumed presence of uncontrolled confounding in addition to selection bias due to fetal death. The observed association of PFAS and ADHD appeared to be either positive or negative in this scenario, and largely depended on the direction and magnitude of the uncontrolled confounding effect; the strongest inverse association between PFAS and ADHD ($OR_{adjusted}=0.73$, bias ratio = 1.37) was estimated with PFAS having a presumed strong impact on fetal loss (OR_{PFAS-C}, $OR_{PFAS-S1}$, $OR_{PFAS-S2}=0.5$) and strong uncontrolled negative confounding bias ($OR_{ADHD-U3}=4$, $OR_{U3-PFAS}=0.3$, prevalence of U3 = 0.2).

In sensitivity analyses, inverse associations between PFAS and ADHD persisted but were smaller in magnitude when PFAS only affected EPL but not conception or clinically observed abortions (Supplementary Figure 2, available as Supplementary data at *IJE* online). The point estimates remained the same but simulation intervals widened when we reduced the size of the simulated cohort (Supplementary Figure 3, available as Supplementary data at *IJE* online), and results changed minimally even when we increased the prevalence of EPL considerably to 50% (Supplementary Table 1, available as Supplementary data at *IJE* online).

Discussion

In this article we illustrated a common bias mechanism where if the exposures of interest reduced success of conception and influenced fetal survival, especially affecting exposed fetuses at higher risk for the outcome, we would expect to find a negative bias in studies that could necessarily only examine live-born children for the outcome of interest. Our simulations that based on information taken from the Danish National Birth Cohort and several hypothetical scenarios suggest the weak inverse associations observed for prenatal PFAS and ADHD in school-aged children can appear even if PFAS do not cause ADHD but have considerable effects on fetal survival. The magnitude of the bias was generally small given the parameter values

Table 4. Simulation results[a] of prenatal PFAS levels and ADHD in a hypothetical live-born birth cohort (scenario 2 assuming a true causal OR = 1.5 of PFAS on ADHD)

$Pr(U=1)$[c]	OR_{ADHD-U}	OR_{C-U}, OR_{S1-U}, OR_{S2-U} =					
		0.8		0.5		0.3	
		Crude OR	Adjusted OR[b]	Crude OR	Adjusted OR[b]	Crude OR	Adjusted OR[b]
where OR_{C-PFAS}, $OR_{S1-PFAS}$, $OR_{S2-PFAS}$ = 0.8							
0.2	2	1.45	1.50 (1.38-1.62)	1.45	1.49 (1.36-1.62)	1.44	1.48 (1.36-1.63)
0.2	5	1.42	1.47 (1.36-1.57)	1.41	1.45 (1.34-1.56)	1.39	1.44 (1.33-1.56)
0.2	10	1.37	1.42 (1.33-1.51)	1.36	1.40 (1.31-1.50)	1.34	1.38 (1.29-1.49)
0.4	2	1.44	1.49 (1.37-1.61)	1.43	1.48 (1.37-1.61)	1.42	1.47 (1.34-1.62)
0.4	5	1.40	1.46 (1.37-1.55)	1.39	1.44 (1.34-1.54)	1.36	1.41 (1.30-1.53)
0.4	10	1.36	1.41 (1.34-1.49)	1.33	1.39 (1.30-1.46)	1.30	1.35 (1.26-1.44)
where OR_{C-PFAS}, $OR_{S1-PFAS}$, $OR_{S2-PFAS}$ = 0.5							
0.2	2	1.37	1.49 (1.36-1.63)	1.34	1.46 (1.33-1.61)	1.34	1.45 (1.30-1.61)
0.2	5	1.33	1.45 (1.34-1.57)	1.27	1.39 (1.27-1.51)	1.25	1.35 (1.22-1.48)
0.2	10	1.28	1.39 (1.30-1.50)	1.21	1.31 (1.21-1.42)	1.16	1.25 (1.14-1.37)
0.4	2	1.36	1.48 (1.35-1.61)	1.32	1.44 (1.31-1.59)	1.30	1.42 (1.26-1.57)
0.4	5	1.31	1.44 (1.34-1.53)	1.24	1.36 (1.25-1.47)	1.18	1.28 (1.16-1.39)
0.4	10	1.26	1.38 (1.31-1.47)	1.18	1.28 (1.20-1.37)	1.09	1.17 (1.08-1.26)

[a]Assume fixed priors for $Pr(R=1)=0.4$, OR_{C-R}, OR_{S1-R}, $OR_{S2-R}=0.5$, $OR_{ADHD-R}=8$.
[b]Adjusted for R (known and measured risk factors that impacted on fetal survival).
[c] Prevalence of unknown or unmeasured risk factors.

Table 5. Simulation results[a] of prenatal PFAS levels and ADHD in a hypothetical live-born birth cohort (scenario 3 assuming no effect of PFAS on ADHD and multiple uncontrolled risk factors of ADHD)

$OR_{PFAS-U2}$	$OR_{ADHD-U2}$	$OR_{PFAS-U3}$	$OR_{ADHD-U3}$	OR_{C-PFAS}, $OR_{S1-PFAS}$, $OR_{S2-PFAS}$ =	
				0.8	0.5
				Adjusted OR[b]	Adjusted OR[b]
1.5	2	1.5	2	1.07 (0.99-1.16)	1.02 (0.93-1.12)
1.5	2	3	2	1.17 (1.09-1.28)	1.12 (1.02-1.22)
1.5	2	0.5	2	0.94 (0.86-1.03)	0.90 (0.81-0.99)
1.5	2	0.3	2	0.90 (0.82-0.99)	0.86 (0.77-0.95)
1.5	2	1.5	4	1.14 (1.05-1.23)	1.09 (1.00-1.19)
1.5	2	3	4	1.39 (1.30-1.49)	1.33 (1.23-1.43)
1.5	2	0.5	4	0.85 (0.78-0.93)	0.82 (0.73-0.89)
1.5	2	0.3	4	0.77 (0.70-0.84)	0.73 (0.67-0.81)

[a]Assume fixed priors for $Pr(R=1)=0.4$, $Pr(U1=1)=0.1$, $Pr(U2=1)=0.2$, $Pr(U3=1)=0.2$, OR_{C-R}, OR_{S1-R}, $OR_{S2-R}=0.5$, OR_{C-U1}, OR_{S1-U1}, $OR_{S2-U1}=0.3$, OR_{C-U2}, OR_{S1-U2}, $OR_{S2-U2}=0.5$, $OR_{ADHD-R}=8$, $OR_{ADHD-U1}=5$, $OR_{ADHD-U2}=2$.
[b]Adjusted for R (known and measured risk factors that impacted on fetal survival).

we specified (bias ratios range from 1.0 to 1.4 for all estimates in scenario 1 and 2). Adjusting for common causes of the outcome that also influence fetal deaths may be a practical way to reduce or eliminate the 'live-birth bias'.

When studying childhood diseases as well as long-term health effects that result from fetal exposure, conditioning on live birth is inevitable since only those who survive are candidates for (at risk of) longer-term health outcomes. Our limited simulation study illustrates that if exposed fetuses are at higher risk of both fetal death (as a competing outcome) and the outcome of interest, either we underestimate the total adverse effect by only considering those who survive till birth, or the exposure may even appear to be protective if the exposure has no effect on the outcome. Conditioning on birth status opens up a 'collider' path and induces bias via other uncontrolled factors that cause both fetal death and the outcome. This bias mechanism can also be conceptualized as a form of competing risk bias, as illustrated previously.[26,37] Studies relating prenatal exposures to long-term health outcomes require reflection

on the influence of competing mortality risks, but many studies at present have to assume no competing risks or that fetal deaths happen at random.

The aetiology of ADHD is not well understood, but both environmental and genetic factors are expected to contribute.[32,33] Several cross-sectional studies have reported a positive correlation between current serum level of PFAS and ADHD in children;[38,39] however, recent longitudinal studies found some unexpected inverse associations between prenatal PFAS levels and ADHD risks in children.[10–12] Our simulation study provided an alternative explanation where PFAS are having an impact on fetal loss and lead to bias suggesting protective associations for prenatal PFAS and ADHD. PFAS may have eliminated some of the fetuses most susceptible to ADHD, thus resulting in lower prevalence of ADHD cases in children through follow-up.

Some caveats should be considered when interpreting our simulation results. We used the PFAS and ADHD example merely to illustrate the magnitude and direction of possible biases induced by conditioning on live-born status in observational research if the assumptions we made here hold. Although prenatal exposure to PFAS was shown to cause fetal and neonatal deaths in animals, it is still debated whether PFAS may have an impact on conception or fetal loss in humans.[20,21,40] Moreover, it is in general reasonable to believe that there are uncontrolled risk factors for ADHD, but knowledge of whether or how strong these factors are correlated with fecundity and fetal loss is limited; we thus varied the strength of the presumed associations for all unknown factors and presented multiple scenarios from simulations. The estimates in our study should also not be directly compared with previous reports based on actual data and models from the DNBC, because we used much simpler scenarios with all variables having a binary response and we also assumed that there are no effect measure modifications and no other forms of bias such as measurement error present.

'Live-birth bias' is unavoidable in life-course epidemiology studies that investigate the impact of prenatal exposures on long-term health outcomes. For the bias mechanisms we proposed, one way to reduce the bias from conditioning on fetal survival in size is to adjust for common causes of the outcome and fetal loss. Researchers should consider adjusting for measured risk factors of the study outcome that potentially can also impact on fecundity and fetal development, such as if the exposures under study were suspected to cause fetal death. This also highlights the needs for improving our understanding of the determinants of pregnancy loss that are currently not well known, and the importance of collecting such data in pregnancy exposure studies. Some of the risk factors for pregnancy loss could possibly cause or be associated with childhood outcomes studied under the fetal programming hypothesis. However, most likely these factors are unknown or difficult to measure, such as genetic factors and environmental chemical factors. Thus quantitative bias and sensitivity analyses should be conducted to evaluate the potential impact and magnitude of this bias on the results.

Supplementary Data

Supplementary data are available at *IJE* online.

Funding

This work was supported by the Danish Strategic Research Council (FETOTOX10-092818). O.A.A. was partly supported by grants R01-DK095668 from NIH/NIDDK, R01-ES010544 from NIEHS/NINDS, R21-ES022391 from NIEHS and R01-HD072296-01A1 from the Eunice Kennedy Schriver National Institute of Child Health and Human Development.

Conflict of interest: None declared.

References

1. Wilcox AJ, Weinberg CR, O'Connor JF *et al*. Incidence of early loss of pregnancy. *N Engl J Med* 1988;**319**:189–94.
2. Chard T. Frequency of implantation and early-pregnancy loss in natural cycles. *Baillieres Clin Obstet Gynaecol* 1991;**5**:179–89.
3. Bonde JPE, Hjollund NHI, Jensen TK *et al*. A follow-up study of environmental and biologic determinants of fertility among 430 Danish first-pregnancy planners: Design and methods. *Reprod Toxicol* 1998;**12**:19–27.
4. Philipp T, Kalousek DK. Generalized abnormal embryonic development in missed abortion: embryoscopic and cytogenetic findings. *Am J Med Genet* 2002;**111**:43–7.
5. Giacomucci E, Bulletti C, Polli V, Prefetto RA, Flamigni C. Immunologically mediated abortion (IMA). *J Steroid Biochem Mol Biol* 1994;**49**:107–21.
6. Weselak M, Arbuckle TE, Walker MC, Krewski D. The influence of the environment and other exogenous agents on spontaneous abortion risk. *J Toxicol Environ Health B Crit Rev* 2008;**11**: 221–41.
7. Kline J, Stein ZA, Susser M, Warburton D. Smoking: a risk factor for spontaneous abortion. *N Engl J Med* 1977;**297**:793–6.
8. Mason CA, Kirby RS, Sever LE, Langlois PH. Prevalence is the preferred measure of frequency of birth defects. *Birth Defects Res A Clin Mol Teratol* 2005;**73** 690–92.
9. Fei C, Olsen J. Prenatal exposure to perfluorinated chemicals and behavioral or coordination problems at age 7 years. *Environ Health Perspect* 2011;**119**:573–78.
10. Liew Z, Ritz B, von Ehrenstein OS *et al*. Attention deficit/hyperactivity disorder and childhood autism in association with prenatal exposure to perfluoroalkyl substances: a nested case-control study in the danish national birth cohort. *Environ Health Perspect* 2014; doi:10.1289/ehp.1408412.

11. Ode A, Kallen K, Gustafsson P *et al*. Fetal exposure to perfluorinated compounds and attention deficit hyperactivity disorder in childhood. *Plos One* 2014;**9**:e95891.
12. Stein CR, Savitz DA, Bellinger DC. Perfluorooctanoate and neuropsychological outcomes in children. *Epidemiology* 2013; **24**:590–99.
13. Johansson N, Fredriksson A, Eriksson P. Neonatal exposure to perfluorooctane sulfonate (PFOS) and perfluorooctanoic acid (PFOA) causes neurobehavioural defects in adult mice. *Neurotoxicology* 2008;**29**:160–69.
14. Johansson N, Eriksson P, Viberg H. Neonatal exposure to PFOS and PFOA in mice results in changes in proteins which are important for neuronal growth and synaptogenesis in the developing brain. *Toxicol Sci* 2009;**108**:412–18.
15. Yahia D, Tsukuba C, Yoshida M, Sato I, Tsuda S. Neonatal death of mice treated with perfluorooctane sulfonate. *J Toxicol Sci* 2008;**33**:219–26.
16. Lau C, Thibodeaux JR, Hanson RG *et al*. Effects of perfluorooctanoic acid exposure during pregnancy in the mouse. *Toxicol Sci* 2006;**90**:510–18.
17. Luebker DJ, York RG, Hansen KJ, Moore JA, Butenhoff JL. Neonatal mortality from in utero exposure to perfluorooctanesulfonate (PFOS) in Sprague-Dawley rats: dose-response, and biochemical and pharamacokinetic parameters. *Toxicology* 2005;**215**:149–69.
18. Wen LL, Lin LY, Su TC, Chen PC, Lin CY. Association between serum perfluorinated chemicals and thyroid function in U.S. adults: the National Health and Nutrition Examination Survey 2007-2010. *J Clin Endocrinol Metab* 2013; **98**:e1456–64.
19. Kjeldsen LS, Bonefeld-Jorgensen EC. Perfluorinated compounds affect the function of sex hormone receptors. *Environ Sci Pollut Res Int* 2013;**20**:8031–44.
20. Fei C, McLaughlin JK, Lipworth L, Olsen J. Maternal levels of perfluorinated chemicals and subfecundity. *Hum Reprod* 2009; **24**:1200–5.
21. Buck Louis GM, Sundaram R, Schisterman EF *et al*. Persistent environmental pollutants and couple fecundity: the LIFE study. *Environ Health Perspect* 2013;**121**:231–6.
22. Darrow LA, Howards PP, Winquist A, Steenland K. PFOA and PFOS serum levels and miscarriage risk. *Epidemiology* 2014;**25**: 505–12.
23. Toft G, Jonsson BAG, Lindh CH *et al*. Exposure to perfluorinated compounds and human semen quality in arctic and European populations. *Hum Reprod* 2012;**27**: 2532–40.
24. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;**10**:37–48.
25. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;**15**:615–25.
26. Thompson CA, Zhang ZF, Arah OA. Competing risk bias to explain the inverse relationship between smoking and malignant melanoma. *Eur J Epidemiol* 2013;**28**:557–67.
27. VanderWeele TJ. Commentary: Resolutions of the birthweight paradox: competing explanations and analytical insights. *Int J Epidemiol* 2014;**43**:1368–73.
28. 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.
29. Russell G, Ford T, Rosenberg R, Kelly S. The association of attention deficit hyperactivity disorder with socioeconomic disadvantage: alternative explanations and evidence. *J Child Psychol Psychiatry* 2014;**55**:436–45.
30. Zhu JL, Olsen J, Liew Z, Li J, Niclasen J, Obel C. Parental smoking during pregnancy and ADHD in children: the Danish national birth cohort. *Pediatrics* 2014;**134**:e382–8.
31. Li J, Olsen J, Vestergaard M, Obel C. Attention-deficit/hyperactivity disorder in the offspring following prenatal maternal bereavement: a nationwide follow-up study in Denmark. *Eur Child Adolesc Psychiatry* 2010;**19**:747–53.
32. Thapar A, Cooper M, Eyre O, Langley K. What have we learnt about the causes of ADHD? *J Child Psychol Psychiatry* 2013;**54**: 3–16.
33. Akutagava-Martins GC, Salatino-Oliveira A, Kieling CC, Rohde LA, Hutz MH. Genetics of attention-deficit/hyperactivity disorder: current findings and future directions. *Expert Rev Neurother* 2013;**13**:435–45.
34. Olsen J, Melbye M, Olsen SF *et al*. The Danish National Birth Cohort – its background, structure and aim. *Scand J Public Health* 2001;**29**:300–7.
35. Norsker FN, Espenhain L, Rogvi SA, Morgen CS, Andersen PK, Andersen AMN. Socioeconomic position and the risk of spontaneous abortion: a study within the Danish National Birth Cohort. *BMJ Open* 2012;**25**;2e001007.
36. Wicklin R. *Simulating Data with SAS*. Cary, NC: SAS Institute, 2013.
37. Andersen PK, Geskus RB, de Witte T, Putter H. Competing risks in epidemiology: possibilities and pitfalls. *Int J Epidemiol* 2012; **41**:861–70.
38. Stein CR, Savitz DA. Serum perfluorinated compound concentration and attention deficit/hyperactivity disorder in children 5-18 years of age. *Environ Health Perspect* 2011;**119**:1466–71.
39. Hoffman K, Webster TF, Weisskopf MG, Weinberg J, Vieira VM. Exposure to polyfluoroalkyl chemicals and attention deficit/hyperactivity disorder in U.S. children 12-15 years of age. *Environ Health Perspect* 2010;**118**:1762–7.
40. Vestergaard S, Nielsen F, Andersson A-M *et al*. Association between perfluorinated compounds and time to pregnancy in a prospective cohort of Danish couples attempting to conceive. *Hum Reprod* 2012;**27**:873–80.