Selection Bias Due to Parity-conditioning in Studies of Time Trends in Fertility

Markku Sallmén,[a] Jens Peter Bonde,[b] Marja-Liisa Lindbohm,[a] and Petter Kristensen[c,d]

Background: Studies of couple fertility over time have often examined study populations with broad age ranges at a cross-section of time. An increase in fertility has been observed in studies that followed episodes of fertility events either prospectively among nulliparous women or retrospectively among parous women. Fertility has a biological effect on parity. If defined at a cross-section of time, parity will also be affected by year of birth, and thus becomes a collider. Conditioning (stratifying, restricting, or adjusting) on a collider may cause selection bias in the studied association.

Methods: A study with prospective follow-up was taken as the model to assess the validity of fertility studies. We demonstrate the potential for selection bias using causal graphs and nationwide birth statistics and other demographic data. We tested the existence of parity-conditioning bias in data including both parous and nulliparous women. We also used a simulation approach to assess the strength of the bias in populations with prior at-risk cycles. Finally, we evaluated the potential for selection bias due to conditioning on parity in various sampling frames.

Results: Analyses indicate that the observed increase in fertility over time can be entirely explained by selection bias due to parity-conditioning.

Conclusion: Heterogeneity in fertility and differential success in prior at-risk cycles are the ultimate factors behind the selection bias. The potential for selection bias due to parity-conditioning varies by sampling frame. A prospective multidecade study with representative sampling of birth cohorts and follow-up from menarche to menopause would bypass the described bias.

(*Epidemiology* 2015;26: 85–90)

Submitted 27 December 2013; accepted 01 July 2014; posted 27 October 2014.
From the [a]Finnish Institute of Occupational Health, Centre of expertise for Health and Work Ability, Helsinki, Finland; [b] Department of Occupational and Environmental Medicine, Bispebjerg Hospital University of Copenhagen, Copenhagen, Denmark; [c]Department of Occupational Medicine and Epidemiology, National Institute of Occupational Health, Oslo, Norway; and [d]Institute of General Practice and Community Medicine, University of Oslo, Oslo, Norway.
The authors report no conflicts of interest.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com). This content is not peer-reviewed or copy-edited; it is the sole responsibility of the authors.
Correspondence: Markku Sallmén, Finnish Institute of Occupational Health, Centre of expertise for Health and Work Ability, Topeliuksenkatu 41 aA, FI-00250, Helsinki, Finland. E-mail: markku.sallmen@ttl.fi.

ISSN: 1044-3983/15/2601-0085
DOI: 10.1097/EDE.0000000000000190

Since the milestone paper by Carlsen et al,[1] potentially decreasing trends in male fertility have been debated worldwide, and the reproductive toxicity of environmental xenobiotics has become a concern in many countries. The few studies of couple fertility have variously shown increasing[2,3] and decreasing[4] fertility over time, or no clear trend[5] (Table 1). None of the studies has followed the reproductive success of birth cohorts from menarche to menopause, or measured the fertility of women at all reproductive ages throughout the entire study period.[2–5] Instead, the study populations have been established at a cross-section of time. In particular, two studies clearly showing an increase in fertility over time conditioned on parity.[2,3] These study decisions motivated us to assess whether the observed increase could be caused by selection bias.[6]

In this paper, we focus on questionable conditioning on parity. As further reading, we refer to earlier studies that have discussed a number of selection problems and other methodologic issues in studies of fertility over time.[3,7–10] A recent study using 5 European data set points out the need to consider cohort and period effects simultaneously.[11] Separate analyses of period and cohort effects in these data showed an increase in fertility over time, but a decrease was observed in the earlier female birth cohorts in mutually adjusted analysis. In eAppendix 1 (http://links.lww.com/EDE/A838), we present selected key concepts and selection processes in studies of time trends in fertility.

STUDIES OF TIME TRENDS IN FERTILITY

Table 1 presents the characteristics of the 4 studies on couple fertility over time.[2–5] Time to pregnancy,[2] subfertility,[3] primary infertility,[4] or their combinations[5] were used as a measure of fertility. Study populations were established at a cross-section of time; 3 studies explored fertility events in the past,[2,4,5] and 1 study, excluding parous women, focused on subsequent pregnancies.[3] Data on fertility events were collected at the first antenatal care visit in 1 study,[3] and 1–51 years after fertility events in other studies.[2,4,5] One study[3] was nationwide and included registered fertility information on all the births. Other studies collected data on fertility by interviews from population samples.[2,4,5] In 3 studies, the age range at the cross-section was wide, including the entire reproductive age,[2–4] but nearly all of the selected women in 1 study

TABLE 1. Studies on Time Trends in Fertility in Populations Established at a Cross-section of Time

Study	Cohort Definition	Cross-section Year; Age Range (years)	Time of Data Collection in Relation to Fertility Events	No. Participants (participation percent)	Focus	Fertility over Time
Notkola (1995)[4]	Random sample from 1938 to 1967 birth cohorts in Finland	1989 22–51	1–29 years later	4,155 (81.4)	All pregnancies and all periods of infertility included	C: decrease
Joffe (2000)[2]	Representative sample of British population aged 16–59 years in 1996	1996 16–59	3–35 years later	1,540 (~70)	1st birth of parous women	P: increase and C: weak increase[a]
Jensen et al (2005)[5]	Danish twins from 1931 to 1952 birth cohorts	1998–1999 45–66	1–51 years later	1,715 women (81.3)	1st attempt	C and P: no change or weak increase
Scheike et al (2008)[3]	Childless Swedish women on 31 December 1982	1982 All ages	At birth	832,000 (very high)	1st birth	C and P: increase

[a]Findings were not represented numerically. "The trend was much weaker in cohort analysis than in period analysis.[2]"
C, cohort effect; P, period effect.

had reached the end of their reproductive career.[5] One study focused solely on the birth cohort effect,[4] whereas 3 studies analyzed both cohort and period effects.[2,3,5]

STUDYING BIRTH COHORT EFFECT: SELECTION BIAS DUE TO CONDITIONING ON PARITY

In general, associations between year of birth, fertility, and parity can be disentangled by a causal diagram (Figure A). The direction of causality is from biological fertility, through fecundability, to parity.[9] Thus, fertility affects parity, and not vice versa. In addition, at a cross-section of time, causality goes from the year of birth to parity, but not vice versa. This is simply because parity increases along with increasing age. Parity is thus a collider, and in line with the graph theory, restricting the study to nulliparous women at a cross-section of time seems to cause a design-driven bias in the association between year of birth and fertility.[6,12]

We used the Swedish study of cohort effects[3] as a model in assessing validity because of the availability of nationwide birth statistics and other demographic data. The study included all nulliparous women living in Sweden on 31 December 1982, who subsequently gave birth to their first child in 1983–2002. Cohort effects were analyzed for women born between 1945 and 1979. We believe that the two conditions—restriction to nulliparous at a cross-section of time and reproduced for the first time in a given period—have caused a selection bias toward increasing fertility over time due to the selective participation of the less fertile women in the earlier birth cohorts, and the more fertile women in the later birth cohorts. This study assumes that the excluded women, who had either reproduced before 1983 at younger ages or reproduced for the first time at an older age after 2002, have fertility similar to that of the eligible women from the same birth cohorts—which seems to be unlikely.

Selection by Year of Birth and Parity

The inclusion and exclusion of women by birth cohort in the Swedish study[3] are described in Table 2. Using Swedish population statistics data on the mean female population in 1984 and 1989,[13] and the numbers of included pregnancies by birth cohorts,[3] we estimated the proportion of women contributing to the follow-up. In the birth cohorts of 1945 to the late 1960s, childlessness increased slightly in Sweden from about 12% to about 14%.[14] Combining the 3 above-mentioned data sources,[3,13,14] we estimated the birth cohort-specific inclusion percentage among women who had ever delivered or would have delivered to be about 3%, 13%, 46%, 86%, 91%, 60%, and 24% for the 7 consecutive birth cohorts of 1945–1949 to 1975–1979, respectively (Table 2). Women, mainly from the earlier birth cohorts, who had their first child at a younger age before 1983, were excluded. In contrast, the study included women, mainly from the later birth cohorts, who delivered their first child at a younger age during the study period (Table 2). Thus, inclusion was strongly related to the year of birth and parity across birth cohorts.[3]

Selection by Fertility

Not only the inclusion percentage but also the underlying reasons behind inclusion differed between the birth cohorts of the study.[3] With the sampling frame of that study, the association between parity and fecundability has previously been studied at cross-sections of time in women's lives; here, the cross-section was the calendar time of beginning to try to conceive. On the basis of several retrospective studies, the age-adjusted fecundability of parous women appears to be about

TABLE 2. Inclusion and Exclusion of Swedish Ever-delivered Women by Year of Birth and Observed Subfertility in the First Calendar Time Period in Scheike et al's study[3]

Birth Cohort	Age in 1983 (years)	Age in 2002 (years)	Size of Birth Cohort in Sweden (No.)	Included in the Study No. (%)	Childlessness (%) After Menopause[13]	Inclusion Percentage Among Ever-delivered Women	Exclusion Percentage Due to Timing of First Child: Before 1983	Exclusion Percentage Due to Timing of First Child: After 2002	Observed Subfertility (%) in the First Calendar Time Period[a]
1945–1949	34–38	53–57	327,588	7,447 (2)	12	~3	~86	0	14.5
1950–1954	29–33	48–52	286,467	32,226 (11)	13	~13	~76	0	15.0
1955–1959	24–28	43–47	275,844	111,200 (40)	13	~46	~47	0	13.8
1960–1964	19–23	38–42	277,831	207,691 (75)	13	~86	~10	~2	8.4
1965–1969	14–18	33–37	288,176	227,576 (79)	13	~91	0	~8	6.7
1970–1974	9–13	28–32	275,663	140,904 (52)	14	~60	0	~34	6.7
1975–1979	4–8	23–27	248,547	51,559 (21)	14	~24	0	~64	4.7

[a]First calendar time period refers to ages below 25 years for cohort categories born later than 1954, 25–29 age group for 1950–1954 birth cohort, and 30–34 age group for 1945–1949 birth cohort.

20–50% higher than that of nulliparous women.[15–18] Moreover, the risk of delayed conception (>1 year) of nulliparous women was more than twice (28%) that of parous women (12%) in a prospective study.[15] Thus, a population of reproductive-aged women, defined at a cross-section of time, has two components: (1) parous women who, on average, are more fertile than their nulliparous counterparts and (2) less fertile nulliparous women. Consequently, a follow-up study that excludes parous women would show fecundability for earlier birth cohorts to be lower than it actually was. In effect, the inclusion was differentially related to fertility, as in the Swedish study.[3] We hypothesize that the magnitude of the bias due to parity-conditioning is related to the birth year-specific proportion of excluded women at the cross-section of time (for earlier birth cohorts) and at the end of follow-up (for later birth cohorts) (Table 2). We further propose that heterogeneity in fertility and differential success in episodes of unprotected intercourse by fertility in the past are the ultimate factors behind the selection bias.

Evidence of Parity-conditioning Bias from Other Data

To empirically test the existence of parity-conditioning bias, we studied the birth cohort effect using data from 4 earlier studies among Finnish families of female and male workers.[15,20–22] The range of the women's years of birth was 1933–1962. One pregnancy from 1973 to 1983 was randomly selected for each woman from hospital records. Data on time to pregnancy and parity were collected using questionnaires. The age of the woman was defined according to the time-to-pregnancy starting date. We considered truncation bias by restricting time-to-pregnancy starting dates between 1 July 1972 and 31 December 1981. The data on time to pregnancy by continuous maternal year of birth were analyzed using discrete proportional hazards regression (SAS PROC PHREG procedure with exact handling of ties) providing fecundability density ratios (FDRs). All analyses were carried out using the Statistical Analysis System 9.4 program (SAS Institute, Cary, NC).

No trend in fertility was seen in the age-adjusted analysis of birth cohort effect, which included all the women (n = 950; FDR = 1.00 [95% confidence interval (CI) = 0.98–1.02]). A subset analysis of nulliparous women (n = 338) induced a strong trend toward an increase in fertility for each increasing year of birth (FDR = 1.04 [1.00–1.07]). Thus, restriction to nulliparous women seems to have caused a selection bias in a setting not particularly designed for studying trends in fertility.

We cannot assess the true trend in fertility on the basis of these data because the original studies were not planned to assess time trends. Moreover, the analysis that includes one period of time to pregnancy per woman, independent of parity, may still suffer from attrition-related bias and from other biases due to, for example, co-linearity problem. The effect of age, period, and cohort cannot be analyzed in the same model because of linear dependence of each other.[11] Nevertheless, the data include both nulliparous and parous women, defined at the time-to-pregnancy initiation, and could therefore be used for demonstration purposes.

In eAppendix 2 (http://links.lww.com/EDE/A838), we describe a simulation approach to assess change in fertility in populations with varying proportions of childless women who have experienced at-risk cycles in the past. Both the number of prior at-risk cycles and proportion of women with prior at-risk cycles affect fertility and thus the strength of selection bias.

Key Issues in Selection Bias Due to Parity-conditioning

Parity-conditioning is a selection process that can manifest itself in a cross-sectional population sample of a wide age range. If such a sample is conditioned on parity, parity-conditioning operates during the reproductive age of the women. As a distinction to truncation bias, in that selection problems

TABLE 3. Parity-conditioning Bias and Attrition-related Bias[a] in Studies of Fertility over Time, by Direction of Follow-up. In all the Designs, Study Population has been Defined at a Cross-section of Time, and the Focus is on Cohort Effects.

Age Range at Cross-section of Time, and Conditioning on Parity	Susceptibility to Parity-Conditioning and Attrition-related Biases	Study
Retrospective follow-up		
Women past menopause or finished reproductive career	Free from both biases	Jensen et al[5]
Wide age range at cross-section of time		
No conditioning on parity	Free from both biases	Notkola[4]
Conditional on parity	Susceptible to parity-conditioning bias Free from attrition-related bias	Joffe[2]
Prospective follow-up		
Cross-section before menarche (multidecade follow-up)	Free from both biases	No available studies
Wide age range at cross-section		
No conditioning on parity	Oldest birth cohorts: susceptible to attrition-related bias Youngest birth cohorts: susceptible to parity-conditioning bias	No available studies
Conditional on parity	Oldest birth cohorts: susceptible to both biases Youngest birth cohorts: susceptible to parity-conditioning bias	Scheike et al.[3]

[a]Together with low desired family size, heterogeneity in fertility leads to a selection process called attrition. Attrition is a process that operates before the study period through differential success of pregnancies among fertile and less fertile couples. The most fertile couples reach the desired family size at a younger age and are not at risk in further follow-up.

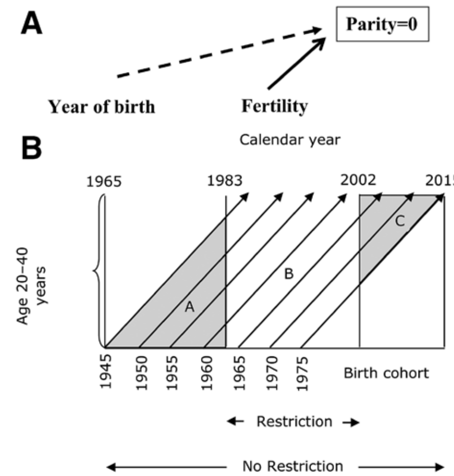

FIGURE. Two simplified illustrations of causal relation between year of birth, fertility, and parity in study populations established at a cross-section of time. Determinants of fertility that may have been changing over time (eg, changes in demographic patterns, smoking, hygiene, and environmental exposures) are not included in the figures. (**A**) A causal diagram. A causal relation between fertility and parity (solid line) is biologic. Once the study is restricted to nulliparous women (conditioning on parity marked with a rectangle), a strong conditional design-driven relation between year of birth and parity is induced (dashed line). Parity is a common effect of year of birth (exposure), and fertility (outcome), ie, a collider. According to graph theory,[6,12] conditioning on a common effect opens the path between year of birth and fertility association, thereby inducing selection bias. (**B**) Lexis diagram for 1945–1975 birth cohorts in a study including women who gave birth during the study period 1983–2002 (area B). The areas refer to the aging of participants and to dates for the initiation of pregnancy attempts at the age of 20 years or older. The grey area A includes 3 types of ever-reproducing women: (a) Women who gave birth to one or more children prior to 1983; and childless women (b) with no, or (c) with one or more prior cycles at risk of pregnancy. The grey area C refers to women who reproduce for the first time after the study period, and are thus excluded. A study without any restriction on age or parity would include first at-risk cycles of women from ~1960 to 2020.

in inclusion lie at both ends of the study period, parity-conditioning can also be viewed as truncation during reproductive age (Figure B). We propose that heterogeneity in fertility and differential success in episodes of unprotected intercourse in the past are the ultimate factors behind the selection bias.

Potential for Selection Bias in Studies of Cohort Effect: Impact of Sampling Frame

In Table 3, we summarize the potential for parity-conditioning bias and attrition-related bias in studies of cohort effects with varying sampling frames. The study among Danish twins[5] included only women who had finished their reproductive career, leaving no room for the proposed bias. The same study found little evidence of change in fertility across birth cohorts. The Finnish study not only included women within the reproductive age range at the time of interview but also requested information on periods of primary infertility.[4] It therefore does not suffer from parity-conditioning bias, but the findings are supposedly biased toward decreased fertility over time due to infertility misclassification.[9] Moreover, the study provides very little information on the fertility of the youngest birth cohorts.

The British study consisted of a representative sample of the population aged 16–59 years in 1996.[2] Time to pregnancy for the first birth during 1961–1993 was included. This study was also conditional on parity, but, in contrast with the Swedish study,[3] it included only parous women at the cross-section of time. This sampling design supposedly did not cause parity-related selection in the oldest birth cohorts, as they had completed their reproductive career. However, the proportion of highly fertile women will gradually increase among the included women along with increasing year of birth. Thus, a study conditional on parity and focusing on past fertility events is liable to be biased toward increasing fertility over time. To our knowledge, no studies have defined the study population at the age of menarche. As shown above, settings not particularly designated for studying trends in fertility are likely to suffer from the proposed biases, as well, if conditioned on parity.

Selection in the Period Analysis

Could the findings on period effect also be biased? The most fertile women from the earliest birth cohorts were underrepresented in the Swedish study, whereas the latest birth cohorts had an excess of the most fertile women.[3] Moreover, the oldest women contributed mainly to the early years of the study period, and the youngest women to the later years. Thus, it is likely that a study based on nulliparous women is also biased toward increasing fertility over time in the period analysis. It is not plausible that a British study population of parous women aged 16–59 years at a cross-section of time would be valid to study period effects for the last 33 years.[2]

CONCLUDING REMARKS

Earlier studies have documented a number of selection problems that should be considered in studies on trends in fertility. We propose that one must also avoid parity-conditioning bias and must consider attrition-related bias in certain sampling designs. Otherwise, a study will reflect the relation between parity and fecundability, rather than changes in fertility across birth cohorts. Restriction to nulliparous women has been recommended for studies on reproductive health in general.[23] However, restricting the study of cohort effect to nulliparous women may produce misleading results, as we have indicated. A subset of nulliparous women could be used in populations with strong restrictions to premarital sex, eg, some Muslim populations.[24] In an ideal study on cohort effects, the study population should be solely defined on the basis of birth cohorts, and reproductive success from several birth cohorts should be followed from menarche to menopause. A study of period effects would prospectively measure the fertility of women at all reproductive ages during the entire study period.

Any change in couple fertility would be a matter of public health importance.[25] Although there are considerable potential biasing factors in studies of time trends in fertility,[7–10] we share the view of some commentary writers that valid surveillance systems should be established to detect such fertility change.[26–28] Due to methodologic limitations in the conducted studies, the question of a possible decline or increase in human capacity to conceive remains unanswered. The analyses we present herein provide insights into how to avoid parity-conditioning and attrition-related biases in such studies.

ACKNOWLEDGMENTS

We thank Miguel Hernán and Donna Baird for their invaluable comments and critical review of the manuscript.

REFERENCES

1. Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. Evidence for decreasing quality of semen during past 50 years. *BMJ*. 1992;305:609–613.
2. Joffe M. Time trends in biological fertility in Britain. *Lancet*. 2000;355:1961–1965.
3. Scheike TH, Rylander L, Carstensen L, et al. Time trends in human fecundability in Sweden. *Epidemiology*. 2008;19:191–196.
4. Notkola I. New information on the prevalence of infertility. *Uutta tietoa hedelmättömyyden yleisyydestä. Suomen Lääkärilehti*. 1995:50:865–870. (In Finnish)
5. Jensen TK, Joffe M, Scheike T, Skytthe A, Gaist D, Christensen K. Time trends in waiting time to pregnancy among Danish twins. *Hum Reprod*. 2005;20:955–964.
6. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615–625.
7. Weinberg CR, Baird DD, Wilcox AJ. Sources of bias in studies of time to pregnancy. *Stat Med*. 1994;13:671–681.
8. Joffe M, Key J, Best N, Keiding N, Scheike T, Jensen TK. Studying time to pregnancy by use of a retrospective design. *Am J Epidemiol*. 2005;162:115–124.
9. Sallmén M, Weinberg CR, Baird DD, Lindbohm ML, Wilcox AJ. Has human fertility declined over time?: Why we may never know. *Epidemiology*. 2005;16:494–499.
10. Key J, Best N, Joffe M, Jensen TK, Keiding N. Methodological issues in analyzing time trends in biologic fertility: Protection bias. *Am J Epidemiol*. 2009;169:285–293.
11. Joffe M, Holmes J, Jensen TK, Keiding N, Best N. Time trends in biological fertility in Western Europe. *Am J Epidemiol*. 2013;178:722–730.
12. Greenland S. Quantifying biases in causal models: Classical confounding vs collider-stratification bias. *Epidemiology*. 2003;14:300–306.
13. Statistics Sweden. Statistical database. Available at http://www.ssd.scb.se/databaser/makro/produkt.asp?produktid=BE0101&lang=2. Accessed February 22, 2013.
14. Persson L, Statistics Sweden. Trend reversal in childlessness in Sweden. In: Conference of European Statisticians, 28–30 April 2010, Lisbon, Portugal.
15. Sallmén M, Lindbohm ML, Kyyrönen P, et al. Reduced fertility among women exposed to organic solvents. *Am J Ind Med*. 1995;27:699–713.
16. Stanton CK, Gray RH. Effects of caffeine consumption on delayed conception. *Am J Epidemiol*. 1995;142:1322–1329.
17. Jensen TK, Scheike T, Keiding N, Schaumburg I, Grandjean P. Selection bias in determining the age dependence of waiting time to pregnancy. *Am J Epidemiol*. 2000;152:565–572.
18. Burdorf A, Brand T, Jaddoe VW, Hofman A, Mackenbach JP, Steegers EA. The effects of work-related maternal risk factors on time to pregnancy, preterm birth and birth weight: The Generation R Study. *Occup Environ Med*. 2011;68:197–204.
19. Zinaman MJ, Clegg ED, Brown CC, O'Connor J, Selevan SG. Estimates of human fertility and pregnancy loss. *Fertil Steril*. 1996;65:503–509.
20. Sallmén M, Anttila A, Lindbohm ML, Kyyrönen P, Taskinen H, Hemminki K. Time to pregnancy among women occupationally exposed to lead. *J Occup Environ Med*. 1995;37:931–934.
21. Sallmén M, Lindbohm ML, Anttila A, et al. Time to pregnancy among the wives of men exposed to organic solvents. *Occup Environ Med*. 1998;55:24–30.
22. Sallmén M, Lindbohm ML, Anttila A, Taskinen H, Hemminki K. Time to pregnancy among the wives of men occupationally exposed to lead. *Epidemiology*. 2000;11:141–147.
23. Olsen J. Options in making use of pregnancy history in planning and analysing studies of reproductive failure. *J Epidemiol Community Health*. 1994;48:171–174.
24. Issa Y, Sallmén M, Nijem K, Bjertness E, Kristensen P. Fecundability among newly married couples in agricultural villages in Palestine: A prospective study. *Hum Reprod*. 2010;25:2132–2138.
25. Wilcox A. *Fertility and Pregnancy: An Epidemiologic Perspective*. New York: Oxford University Press; 2010:138.

26. Olsen J, Rachootin P. Invited commentary: Monitoring fecundity over time–if we do it, then let's do it right. *Am J Epidemiol*. 2003;157:94–97.
27. Bonde JP, Joffe M, Sallmén M, et al. Validity issues relating to time-to-pregnancy studies of fertility. *Epidemiology*. 2006;17:347–349.
28. te Velde E, Burdorf A, Nieschlag E, et al. Is human fecundity declining in Western countries? *Hum Reprod*. 2010;25:1348–1353.
29. Nguyen RH, Wilcox AJ. Terms in reproductive and perinatal epidemiology: 2. Perinatal terms. *J Epidemiol Community Health*. 2005;59:1019–1021.
30. Weinberg CR, Gladen BC. The beta-geometric distribution applied to comparative fecundability studies. *Biometrics*. 1986;42:547–560.
31. Weinberg CR, Wilcox AJ. Methodologic issues in reproductive epidemiology. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:621.
32. Joffe M, Key J, Best N, Jensen TK, Keiding N. The role of biological fertility in predicting family size. *Hum Reprod*. 2009;24:1999–2006.
33. Porta M, ed. *Dictionary of Epidemiology. Fifth Edition.* New York: Oxford University Press; 2008
34. Bongaarts J. A method for the estimation of fecundability. *Demography*. 1975;12:645–660.

Appendix 1. Key concepts and selection processes in studies of time trends in fertility.

Fertility	We use "fertility" as a synonym of fecundity i.e. reproductive capacity. Fertility is a continuum rather than a dichotomy.
Parity	Number of births for a given woman. [29] In this paper we are mainly interested in parity at a given calendar time. A woman with no previous births is nulliparous; otherwise she is parous.
Fecundability	The probability of conception in a cycle. Fertility and fecundability vary considerably in humans. => heterogeneity
Heterogeneity	Heterogeneity is important in fertility studies because women can have multiple periods of being at risk of pregnancy during their reproductive age. [30, 31] Fertile couples, on average, have more children than their less fertile counterparts. [32]
Attrition	Together with low desired family size, heterogeneity leads to a selection process called attrition. Attrition is a process that operates before the study period through differential success of pregnancies among fertile and less fertile couples. The most fertile couples reach the desired family size at younger age, and are not at risk in further follow-up.
Truncation	Truncation is design-driven selection and operates within the study period by differential inclusion of pregnancies by the length of pregnancy attempt at both ends of the study period. [3] In a cross-sectional population sample, couples with long waiting times are overrepresented at the start of follow-up but underrepresented at the end.
Collider	A variable is called a collider when it is the common effect of two or more variables in a causal path. [33] Conditioning (stratifying, restricting or adjusting) on a collider may cause selection bias in the studied association. [6, 12]

Appendix 2. A simulation approach to assess the strength of bias due to restricting the study to nulliparous women at a cross-section of time

The simulations below refer to a study population defined at a cross-section of time (T0), and to birth cohorts which, at T0, have reached reproductive age, i.e. women aged 20–40+ years in the Swedish study. [3] The first time to pregnancy event after T0, leading to the birth of a child, is followed. The simulations approximately correspond to 1945–1960 birth cohorts, but not to 1961–1979 birth cohorts. [3]

Assumptions:

- An original population (A) of 100,000 women, referring to an entire birth cohort with a mean effective fecundability [34] of 0.3 and STD of 0.188. The women have no prior at risk cycles at T0.
- Four other populations of 100,000 women (B1–B4) with an identical fecundability distribution to that in population A. The women in populations B1–B4 have experienced N = 1, 2, 4, or 8 at risk cycles prior to T0, and, in part, are parous.
- All the women in population A are childless and eligible for the study at T0, but as in the Swedish study [3], only childless women at T0 are eligible for the study in the B1–B4 populations.
- Female age and other determinants of fertility are assumed to be considered adequately. Also, for simplicity, all the pregnancies are assumed to end in birth. We refer here to effective fecundability [34]: pregnancy is defined to have started in a given menstrual cycle if and only if the pregnancy ended in the birth of one or more children.

A simulation study on time to pregnancy was conducted among nulliparous women in populations A and B1–B4. Difference in fecundability or infertility in populations B1–B4 as compared to population A describe the strength of selection bias between women with and without prior at risk cycles.

Findings

The study populations include all the 100,000 women from population A, and a varying number of nulliparous women from populations B1–B4 (Table 1.). The number of eligible population B members depends on the number of prior cycles and fecundability of subjects, and is 70,192 for N=1 prior at risk cycle, 52,495 for N=2, 33,764 for N=4, and 18,105 for N=8. This means that 30% (as could be expected on the basis of a mean fecundability of 0.3), 47.5%, 66.2%, and 81.9% of the women had become pregnant by cycle 1, 2, 4, or 8, respectively.

Table 1. Infertility and fecundability of nulliparous women in population B1–B4 by number of prior at risk cycles not leading to pregnancy as compared to the entire non-selected population A. Number of participants in the reference population (original population A) is 100,000 in all the analyses.

Cycles at risk prior to T0	N of nulliparous women included in populations B1–B4	Subfertility OR (95% CI)	Fecundability density ratio (95% CI)
1	B1: 70,192 (70.2%)	1.37 (1.33–1.41)	0.86 (0.85–0.87)
2	B2: 52,495 (52.5%)	1.70 (1.65–1.75)	0.75 (0.75–0.76)
4	B3: 33,764 (33.7%)	2.49 (2.41–2.58)	0.61 (0.61–0.62)
8	B4: 18,105 (18.1%)	4.09 (3.95–4.25)	0.46 (0.45–0.47)

As expected, fecundability decreases and subfertility increases along with the number of prior at risk cycles.

The pregnant women at T0 in this simulation experiment cannot be directly compared to the eligible women in the Swedish study, [3] rather the eligible women in the Swedish study constitute an unknown proportion of women both without (population A) and with one or more prior at risk cycles (populations B1–B4), who eventually reproduced during the study period.

Next, we aim to mimic what could happen in true human populations over time. The focus is strictly on heterogeneity in fertility and the number of prior unprotected cycles. In true human populations, it is an unrealistic assumption that: A) only women without prior at risk cycles would try to conceive in a given time period, or B) only women with prior at risk cycles would try to conceive in a given time period after T0. As regards parity conditioning, the validity of the Swedish study [3] depends on assumption A.

To conduct a comprehensive simulation to assess the potential strength of parity conditioning, one would need at least the following age and cohort-specific assumptions:

- Prior to T0: Number (distribution) of cycles until women stop trying if not pregnant. This assumption should be made for all ages and cohorts separately.
- N1 and N2: the number of nulliparous women with and without prior at risk cycles (population B above), and the proportion of these women attempting pregnancy during the study period or having unattended at risk cycles during the study period. Data on time to pregnancy among those giving birth to a child. Even better would be collecting information on all at risk cycles independent on attempt. This assumption should also be made for all ages and cohorts separately.

There is no information on non-successful at risk cycles in true human populations. Also, it is unrealistic to assess the number of women in these two groups (N1 and N2) for any birth cohort. However, it is a more reasonable assumption that the N1 / N2 ratio increases along with increasing age, and is thus higher in earlier birth cohorts than in later birth cohorts. In one woman, the number of unprotected cycles can increase over time or remain the same but cannot decline. Should there be a systematic cohort specific change in the N1 / N2 ratio i.e. in proportions of “Women with no prior at risk cycles“ and “Women with one or more prior at risk cycles” attempting pregnancy, a change in fertility over time would be observed.

In Table 2, we show fecundability and subfertility (no pregnancy within 12 cycles) in a) populations of nulliparous women having experienced: 1, 2, 4, or 8 prior cycles, and b) a varying proportion of women with prior cycles as compared to the original population. For each cell in Table 2, the size of the selected population B is the same as that in Table 1. The size of the non-selected population A is set to a given number to achieve the proportions of 20%, 40%, 60% and 80 % for women with prior at risk cycles.

Interpretation

For example, a subfertility per cent of 14.3 (prior cycles=2, proportion=40%) would mean a subfertility odds ratio of about 1.26 in comparison to the original population. This finding is comparable with the findings of the Swedish study [3] for 1945 to 1960 birth cohorts. Relative subfertility decreased from about 1.2 to 1.0 from the 1945 birth cohort to the birth cohort 1960 (Figure 2 in reference [3]).

Simulation findings suggest that restriction to nulliparous women may cause a bias towards increase in fertility over time. The strength of the bias may well equal or even exceed the strength of the observed trend in true studies. [3, 2]

eTable 2. Mean fecundability and subfertility (%) in populations according to the varying proportion of childless women with prior at risk cycles. Original population: Mean fecundability 0.300, subfertile 11.8%.

	Proportion of childless women with prior at risk cycles among those who try to get pregnant.			
Prior at risk cycles	20%	40%	60%	80%
1	Mean fecundability: 0.290 Subfertile: 12.4%	Mean fecundability: 0.280 Subfertile: 13.0%	Mean fecundability: 0.269 Subfertile: 13.8%	Mean fecundability: 0.259 Subfertile: 14.4%
2	Mean fecundability: 0.283 Subfertile: 12.9%	Mean fecundability: 0.266 Subfertile: 14.3%	Mean fecundability: 0.249 Subfertile: 15.6%	Mean fecundability: 0.232 Subfertile: 17.0%
4	Mean fecundability: 0.274 Subfertile: 14.1%	Mean fecundability: 0.247 Subfertile: 16.6%	Mean fecundability: 0.220 Subfertile: 19.3%	Mean fecundability: 0.194 Subfertile: 21.7%
8	Mean fecundability: 0.263 Subfertile: 16.2%	Mean fecundability: 0.225 Subfertile: 20.9%	Mean fecundability: 0.188 Subfertile: 25.6%	Mean fecundability: 0.151 Subfertile: 30.1%

REFERENCES

1. Gilbert-Diamond D, Cottingham KL, Gruber JF, et al. Rice consumption contributes to arsenic exposure in US women. *Proc Natl Acad Sci USA*. 2011;108:20656–20660.
2. Melkonian S, Argos M, Hall MN, et al. Urinary and dietary analysis of 18,470 Bangladeshis reveal a correlation of rice consumption with arsenic exposure and toxicity. *PLoS One*. 2013;8:e80691.
3. Meharg AA, Lombi E, Williams PN, et al. Speciation and localization of arsenic in white and brown rice grains. *Environ Sci Technol*. 2008;42:1051–1057.
4. Jones MR, Tellez-Plaza M, Sharrett AR, Guallar E, Navas-Acien A. Urine arsenic and hypertension in US adults: the 2003-2008 National Health and Nutrition Examination Survey. *Epidemiology*. 2011;22:153–161.
5. National Health and Nutrition Examination Survey: MEC in-person dietary interviewers procedures manual. Available at: http://www.cdc.gov/nchs/data/nhanes/nhanes_03_04/DIETARY_MEC.pdf. US Department of Health and Human Services National Center for Health Statistics, 2002. Accessed July 19, 2012.
6. Sun GX, Williams PN, Carey AM, et al. Inorganic arsenic in rice bran and its products are an order of magnitude higher than in bulk grain. *Environ Sci Technol*. 2008;42:7542–7546.
7. Jackson BP, Taylor VF, Karagas MR, Punshon T, Cottingham KL. Arsenic, organic foods, and brown rice syrup. *Environ Health Perspect*. 2012;120:623–626.
8. Ren X-L, Liu Q-L, Wu D-X, et al. Variations in concentration and distribution of health-related elements affected by environmental and genotypic differences in rice grains. *Rice Sci*. 2006;13:170–178.

Selection Bias Due to Parity-conditioning in Studies of Time Trends in Fertility

To the Editor:

The recent article by Sallmén et al.[1] raises the issue of "parity-conditioning bias": women (or couples) of low biological fertility tend to have, on average, fewer children; their family size is limited by biological factors.[1,2] As the article correctly says, "fertility affects parity," potentially causing bias—but it then gets the bias the wrong way round.

Time to pregnancy can be assessed only for women who have had at least one birth—unless augmented by information on infertile phases, as recommended.[3] To obtain unbiased estimates of fertility in a population that is representative of all women who have given birth, the population of first births is appropriate. This is because *all women who have a birth have a first birth*.

Another way of putting this is that biological fertility is a property of the individual (or couple), not of each pregnancy. Inclusion of more than one pregnancy per woman can therefore introduce bias.

Three of the studies cited by Sallmén et al. studied the first birth or the first attempt at conception. It is just plain wrong to describe these as having "conditioned on parity."

The other study[4] included all pregnancies, so excessive weighting was given to the relatively fertile women, who were counted twice (or more). That article's finding that fertility decreased over time is therefore subject to the bias that as family size decreased over time in Finland (as elsewhere), this over-weighting became less pronounced, generating an apparent—but false—downward trend in biological fertility. This is the one empirical article in the literature where parity-conditioning bias is present.

In the simulation presented by Sallmén et al.,[1] they do not make this particular error, but they make a related one, albeit weaker. They consider only one pregnancy per woman, which is valid. The problem is that the pregnancy can be anywhere in the birth order, randomly chosen. When this is from a second or subsequent pregnancy, it is likely to have been to a relatively fertile woman for the reasons already given. The valid comparison would be confined to first pregnancies; their inclusion of subsequent ones introduces a weaker form of parity-conditioning bias than in the Notkola article.

It should be apparent from this discussion that maternal age does not enter into the consideration of this particular bias. Introducing it, and the causal diagram in Figure A of the article by Sallmén et al., only confuses the situation.

ISSN: 1044-3983/15/266-0e67
DOI: 10.1097/EDE.0000000000000383

Michael Joffe
Department of Epidemiology and Biostatistics
Imperial College London
St. Mary's Campus
London, United Kingdom
m.joffe@imperial.ac.uk

REFERENCES

1. Sallmén M, Bonde JP, Lindbohm ML, Kristensen P. Selection bias due to parity-conditioning in studies of time trends in fertility. *Epidemiology*. 2015;26:85–90.
2. Joffe M, Key J, Best N, Jensen TK, Keiding N. The role of biological fertility in predicting family size. *Hum Reprod*. 2009;24:1999–2006.
3. Joffe M, Key J, Best N, Keiding N, Scheike T, Jensen TK. Studying time to pregnancy by use of a retrospective design. *Am J Epidemiol*. 2005;162:115–124.
4. Notkola I. New information on the prevalence of infertility. *Uutta tietoa hedelmättömyyden yleisyydestä. Suomen Lääkärilehti* 1995;50:865–870 [in Finnish].

The Authors Respond

To the Editor:

We thank Michael Joffe for his letter[1] pertaining to our article on selection bias due to parity conditioning in studies of time trends in fertility.[2] We agree that studying the first pregnancy or attempt in a population that is representative of all women is appropriate, and were happy to see that Michael Joffe shared our view on the direction of causality from fertility to parity. However, in our opinion, he makes incorrect conclusions regarding the existence and direction of the selection bias[2] in the referenced studies.[3–6] Our arguments are as follows:

First, and most importantly, Michael Joffe ignores the key issue behind parity-conditioning bias in this context, namely the definition of a study population with a wide age range at a cross section of time. The Swedish study consisted of all nulliparous women living in Sweden on December 31, 1982, who subsequently gave birth to their first child in 1983–2002.[4] The two

The authors report no conflicts of interest.

ISSN: 1044-3983/15/266-0e67
DOI: 10.1097/EDE.0000000000000384

conditions (nulliparous on December 31, 1982 and reproduced in 1983–2002) resulted in a major difference in eligibility between different birth cohorts, with inclusion varying from 3% (birth cohort 1945–1949) to 91% (1965–1969), and further to 24% (1975–1979).[4] Moreover, eligibility was strongly related to parity at the cross section, and thus also to fertility—a finding that is indicative of selection bias. Women with first birth at a younger age were excluded from the earlier birth cohorts, but included in the later birth cohorts. The opposite was true for women with first birth at an older age. This selection by parity is illustrated in the Lexis diagram (Figure). The most extreme difference in eligibility was between women born in 1945 and 1979. The 1945 cohort members should have had their first child at the age of 37 years or over to be included, whereas the 1979 cohort members should have reproduced before 24 years old. This comparison is simply not reasonable. The ultimate factors behind the selection bias are heterogeneity in fertility and differential success in at-risk cycles before the study period (earlier birth cohorts) or during the study period (later birth cohorts).

Second, the Finnish study[5] did not include more than one pregnancy per woman. The study addressed primary infertility defined as an attempt to become pregnant that lasted 12 months or more before the first birth. Five-year age group-specific primary infertility increased from that of the 1938–1949 birth cohort through that of the 1950–1959 cohort to that of the 1960–1967 cohort. As indicated earlier, the findings of that study[5] may be biased toward decreased fertility over time due to infertility misclassification.[7]

Instead of only including births that occurred during a common, arbitrary study period independent of the year of birth, a follow-up of all the birth cohorts should start at menarche and last until first pregnancy or until menopause. Otherwise the findings may suffer from parity-conditioning bias. In conclusion, as regards the existence and direction of parity-conditioning bias, we firmly maintain our original stand.[2]

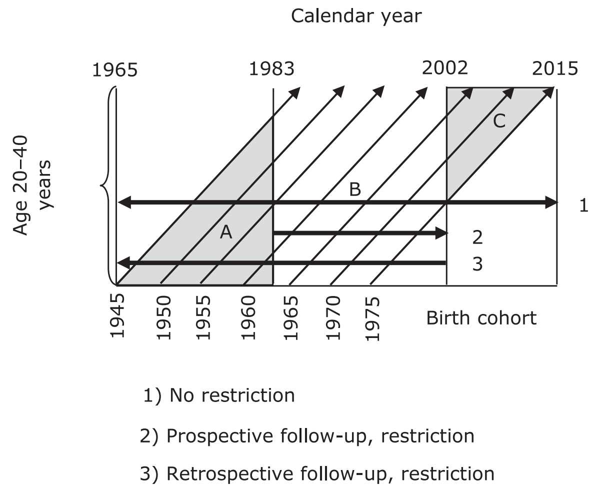

FIGURE. Lexis diagram for 1945–1975 birth cohorts and the follow-up periods of three example studies. *Arrow 1* A study without any restriction on age or parity includes first at-risk cycles of women during ~1960–2020. Two other studies examine study populations with broad age ranges at a cross section of time and condition on parity. *Arrow 2* A study with a prospective follow-up including only women who gave birth during the study period 1983–2002 (area B)[4] excludes women who were parous on 31.12.1982 (gray area A), and those who remained nulliparous after the study period (gray area C).[4] A part of the included women from earlier birth cohorts 1945–1960 supposedly have experienced unsuccessful prior at-risk cycles, and thus are less fertile than those who reproduced before 1983. By contrast, some of the excluded women in the later birth cohorts ~1962–1975 supposedly have experienced unsuccessful at-risk cycles during the study period, and thus are less fertile than those who reproduced. *Arrow 3* A study with a retrospective follow-up excluding women who remain nulliparous at the cross section[3] suffers from the latter selection problem unless information on prior infertile phases or at-risk cycles have been collected.

Markku Sallmén
Finnish Institute of Occupational Health
Centre of Expertise for Health and Work Ability
Helsinki, Finland
markku.sallmen@ttl.fi

Jens Peter Bonde
Department of Occupational and Environmental Medicine
Bispebjerg Hospital University of Copenhagen
Copenhagen, Denmark

Marja-Liisa Lindbohm
Finnish Institute of Occupational Health
Centre of Expertise for Health and Work Ability
Helsinki, Finland

Petter Kristensen
Department of Occupational Medicine and Epidemiology
National Institute of Occupational Health
Oslo, Norway
Institute of Health and Society
University of Oslo
Oslo, Norway

REFERENCES

1. Joffe M. Selection bias due to parity-conditioning in studies of time trends in fertility. *Epidemiology*. 2015:26:e67.
2. Sallmén M, Bonde JP, Lindbohm ML, Kristensen P. Selection bias due to parity-conditioning in studies of time trends in fertility. *Epidemiology*. 2015;26:85–90.
3. Joffe M. Time trends in biological fertility in Britain. *Lancet*. 2000;355:1961–1965.
4. Scheike TH, Rylander L, Carstensen L, et al. Time trends in human fecundability in Sweden. *Epidemiology*. 2008;19:191–196.
5. Notkola I. New information on the prevalence of infertility. *Uutta tietoa hedelmättömyyden yleisyydestä.* Suomen Lääkärilehti. 1995;50:865–870 [in Finnish].
6. Jensen TK, Joffe M, Scheike T, Skytthe A, Gaist D, Christensen K. Time trends in waiting time to pregnancy among Danish twins. *Hum Reprod*. 2005;20:955–964.
7. Sallmén M, Weinberg CR, Baird DD, Lindbohm ML, Wilcox AJ. Has human fertility declined over time?: why we may never know. *Epidemiology*. 2005;16:494–499.