



Published in final edited form as:
J Reprod Med. 2011 ; 56(0): 130–137.

Validity of retrospectively reported behaviors during the periconception window

Audra L Gollenberg, PhD¹, Sunni L Mumford, PhD¹, Maureen A Cooney, PhD¹, Rajeshwari Sundaram, PhD¹, and Germaine M Buck Louis, PhD¹

¹Division of Epidemiology, Statistics and Prevention Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, NIH, DHHS, Bethesda, MD

Abstract

Objective—To assess the validity of retrospectively reported maternal behaviors while attempting pregnancy.

Study Design—Participants in a prospective pregnancy cohort study with periconception enrollment were queried about use of cigarettes, alcohol, vitamins, caffeine, and consumption of sport fish while attempting pregnancy. Prospective longitudinal data reported in daily diaries (gold standard) was compared with data obtained a decade later using a self-administered questionnaire. Agreement was assessed by percent agreement and Kappa coefficients.

Results—Among the 82 participating women, percent agreement ranged from 54 to 74% for the five behaviors. Validity was highest for smoking (Kappa=0.43, 95% Confidence Interval (CI): 0.22, 0.65) followed by fish consumption (Kappa=0.32, 95% CI: 0.09, 0.55), caffeine (Kappa=0.21, 95% CI: 0.09, 0.51), and alcohol (Kappa=0.20, 95% CI: 0.08, 0.33). There were no systematic differences in agreement by time to pregnancy or pregnancy outcome. Associations between smoking and alcohol consumption and pregnancy outcomes were highly sensitive to the levels of misclassification observed in this study.

Conclusion—Validity was poor to moderate for the five behaviors, though higher for more regular behaviors such as smoking and caffeine consumption. The potential for misreporting of periconception behaviors can affect inferences, and thus efforts to capture information prospectively should be promoted.

Keywords

behavior; critical windows; lifestyle; misclassification; periconception; pregnancy; validity

Corresponding author and individual responsible for reprint requests: Dr Audra L Gollenberg, Epidemiology Branch, Division of Epidemiology, Statistics, and Prevention Research *Eunice Kennedy Shriver* National Institute of Child Health and Human Development 6100 Executive Boulevard, Room 7B03, Rockville, MD 20892, Phone: 301-435-6935, Fax: 301-402-2084, algollenberg@gmail.com.

Parts of this work were presented as a poster presentation/discussion at the Annual Meeting of the Society for Pediatric and Perinatal Epidemiologic Research in June 2009.

Introduction

The periconception window is recognized as an important sensitive window for human development,¹ underscoring the importance of preconception guidance for women at risk for (or planning) pregnancy. Such guidance articulates the need for couples, with particular emphasis on women, to adopt a healthy lifestyle (e.g., maintain a healthy weight, abstain from smoking and alcohol, eat a balanced diet, and exercise regularly) to maximize their chances of conceiving and ensuring a healthy pregnancy.^{2, 3} Past authors have reported that women planning a pregnancy change their behaviors in a positive manner during the preconception period.^{4–6} In addition, women with planned pregnancies are reported to have better pregnancy outcomes than women reporting unplanned pregnancies,^{7–9} though the term “planned pregnancy” lacks reliability and is subject to differences in interpretation.^{10, 11} Maternal behaviors during the periconception period have been linked with pregnancy outcomes as well as later onset adult disease.^{12–14} Thus, much emphasis has been placed on forming appropriate preconception guidance for women in order to ensure healthy pregnancy outcomes.^{15, 16} However, information used to influence preconception guidance may be prone to error, as much of the guidance is based on self-reported behaviors. While there has been some attempt to validate retrospectively reported information during pregnancy,^{17–24} to our knowledge there has been no attempt to empirically evaluate the validity of retrospective reporting during the periconception window.

The objective of this study was to assess the long-term validity of retrospectively reported preconception behaviors in relation to prospectively collected information, as well as determine to what degree misclassification of preconception behaviors may affect inferences. We focused on behaviors reported to affect conception probabilities and pregnancy outcomes (e.g., use of cigarettes, alcohol, vitamins, caffeine) as addressed by the Centers for Disease Control and Prevention preconception guidance,²⁵ and also fish consumption given concerns about its use during pregnancy. A second aim was to identify sociodemographic and reproductive factors that affect the validity of reporting.

Methods

Prospective measurement of lifestyle behaviors

The referent study cohort comprised women planning pregnancies who had previously participated in the New York State Angler Cohort Study (n=99), designed to characterize contaminated fish consumption and its association with reproductive outcomes among couples residing in 16 counties surrounding Lakes Erie and Ontario.²⁶ Women were recruited into the prospective pregnancy cohort study in 1996–1998²⁷ and asked to complete daily diaries on sexual intercourse, menstruation, pregnancy test results, and use of cigarettes, alcoholic and caffeinated beverages, consumption of fish, and multivitamin usage daily until pregnant (i.e., two positive home pregnancy tests) or up to 12 at-risk menstrual cycles. Diary compliance was excellent (95%) and considered the gold standard. An at-risk cycle was empirically defined as sexual intercourse during the estimated fertile window, which was defined as an eight-day window ranging from 5 days before through 2 days following ovulation using the Ogino-Knaus method,^{28, 29} given that we did not have a biomarker for ovulation. As a result, time to pregnancy (TTP) was defined as the number of

at-risk cycles required for conception. Full Institutional Review Board Approval was given for the conduct of this study; informed consent was obtained from all women prior to participation.

Retrospective measurement of lifestyle behaviors

In 2006, letters on the study's original letterhead were sent to all women in the study cohort (n=99) irrespective of pregnancy outcome informing them of their serum polychlorinated biphenyl concentrations, and asking for their continued participation in a study focusing on how well women remember what they were doing while trying to conceive. For women with pregnancies resulting in live births, we provided the infant's date of birth and a copy of his/her photograph that was voluntarily sent to the research office by the mother to ensure the mother was reporting for the index pregnancy. Letters were appropriately tailored for women who did not become pregnant or give birth. The follow-up survey asked women if they smoked cigarettes, consumed Great Lakes fish, alcoholic or caffeinated beverages, or took vitamins while trying to become pregnant using a yes/no response. If a woman checked "yes", she was asked to specify the average number while attempting to become pregnant (with the exception of vitamin use). The follow-up survey was completed by 89 women (90% response).

Statistical analysis

A woman was considered to have engaged (or not) in any of the lifestyle behaviors while attempting pregnancy based upon the daily diary data. We estimated the validity of retrospective reporting by estimating the Kappa statistic and accompanying 95% confidence interval (CI), as well as percent agreement, for each lifestyle behavior based on whether the woman reported that she engaged in the behaviors (yes/no) while trying to conceive. We used varying definitions of whether or not the woman engaged in a certain behavior to assess the sensitivity of our measure, since sporadic behavior may be recalled differently than regular behavior. In particular, as a part of our sensitivity analyses, we expanded our definitions of nonuse and then estimated agreement as follows: <5 total cigarettes smoked throughout the entire periconception period, <3 total alcoholic beverages consumed, <5 total caffeinated drinks consumed, <5 total vitamins taken, or <3 total Great Lakes fish meals eaten. Also, we stratified our findings by pregnancy outcome (live birth, pregnancy loss, no conception) and TTP (≤ 6 cycles and >6 cycles for those who conceived). All observed cycles were used for the main analysis, rather than restricting to at-risk cycles, for consistency with the retrospective questionnaire which asked women to recall lifestyle behaviors during the entire periconception period without regard to number of at-risk cycles.

We used unconditional logistic regression to estimate unadjusted odds ratios (ORs) and 95% CIs to identify factors associated with agreement in reporting from among the following covariates: age (in years) upon entry into the prospective pregnancy study, education (college education, yes/no), TTP (in cycles), number of previous pregnancies (0,1, 2+), number of previous live births (0,1, 2+), and pregnancy outcome (live birth, pregnancy loss, and no conception). Covariates were selected *a priori* for their known or suspected association with fecundity and/or reporting of behaviors. Given our limited sample size we did not estimate adjusted ORs in part given the absence of observed associations between

covariates and likelihood of agreement. The main analysis included all women participating in the follow-up study irrespective of their length in the original cohort study or pregnancy outcome. However, only women who completed the study were included in the stratified analyses, given that their entire trying intervals and pregnancy outcomes were observed (n=71).

Spearman correlation coefficients and associated p-values were calculated for daily (prospective) and average (retrospective) usage of cigarettes, alcohol and caffeinated beverages, and Great Lakes fish consumption during the periconception window. Coefficients could not be estimated for vitamin usage, since quantity of vitamin usage was not asked in the follow-up questionnaire. We first estimated correlation coefficients for all participants, then restricted to exclude abstainers (i.e., women whose diaries reflected no usage).

Lastly, we estimated the effect of maternal smoking on live birth using both the retrospective and prospective reports for smoking status to determine the degree to which reporting errors may affect inferences. We then applied the levels of misclassification observed in this study to a hypothetical population of 500 women with varying associations between smoking and live birth (OR=0.8, 1.0, and 1.5) to assess the effects of misclassification among a larger simulated population. All data were analyzed using SAS Version 9.1 (SAS Institute, Cary, NC, USA).

Results

Of the 99 women who participated in the original prospective pregnancy cohort study, 89 (90% response) participated in the follow-up survey of which complete diary data were available for 82 women. Among the 10 non-participating women, two did not achieve pregnancy during the study, one experienced a pregnancy loss, three had live births, and four withdrew before achieving pregnancy. The high response rate for the follow-up survey approximately a decade later may reflect the characteristics of the original cohort, viz., largely white married women with 65% having a college education, and/or the vesting the women had in the original study, given the intensity of daily participation possibly for a year or more.

The overall agreement between the prospectively and retrospectively reported behaviors was fair to moderate and varied by type of behavior (Table I). Reporting errors were bidirectional meaning that women both over- and under-reported behaviors at follow-up. Highest validity was observed for cigarette smoking (Kappa=0.43, 95% CI: 0.22, 0.65), followed by fish consumption (Kappa=0.32, 95% CI: 0.09, 0.55), caffeine consumption (Kappa = 0.21, 95% CI: -0.09, 0.51), and alcohol consumption (Kappa=0.20, 95% CI: 0.08, 0.33). Kappa could not be calculated for vitamin use as noted above. Similar results were noted when comparing percent agreement across the five behaviors, although the order varied slightly: caffeine (85% agreement), cigarette smoking (81%), fish consumption (75%), vitamin use (74%), and alcohol (54%).

Percent agreement increased by approximately 2–3% for each behavior when we expanded our definition of nonuse; Kappa statistics increased slightly for alcohol consumption and cigarette smoking and more than doubled for caffeine consumption (Kappa=0.21 to Kappa=0.47). The Kappa statistic could not be calculated for vitamin use, since all women reported having taken vitamins in the prospective capture of data. However, the Kappa statistic for vitamin use went from not calculable using the strict definition of nonuse to 0.41 using the expanded definition.

In the stratified analyses, Kappa coefficients and percent agreement did not significantly vary across TTP and pregnancy outcome categories (Table II), given overlap of 95% CIs. No clear pattern was observed between percent agreements and TTP or pregnancy outcome. Highest agreement was for caffeine consumption (88%) among women with live births, but cigarette smoking (92%) for women with pregnancy losses and vitamin usage (100%) and caffeine and alcohol (89%) for women without conceptions.

Correlation coefficients for the prospectively and retrospectively reported quantities of each behavior varied by type of behavior and were highest for cigarette smoking ($r=0.71$) followed by caffeine consumption ($r=0.54$), fish consumption ($r=0.41$), and alcohol consumption ($r=0.39$) (Table III). Similar findings were observed when abstaining women (those reporting zeros in daily diaries for the behaviors) were excluded from the analysis.

Reporting errors were bidirectional meaning that women both over- and under-reported behaviors at follow-up. Moreover, no systematic differences were observed for any covariates, nor were there differences by overall reporting agreement (yes/no) across the five behaviors (data not shown). However, higher parity was associated with a significant 90% reduction in the likelihood of agreement for alcohol consumption (OR=0.10, 95% CI: 0.03, 0.38), and an 82% significant reduction for vitamin usage (OR=0.18, 95% CI: 0.04, 0.93). Similarly, TTP >6 months was associated with an 89% significant reduction in the likelihood of agreement for cigarette smoking (OR=0.11, 95% CI: 0.02, 0.59).

We observed that ORs for the association between retrospective smoking data and live birth were highly sensitive to misclassification. Although confidence intervals were wide, ORs based on retrospective reporting suggested a positive association between smoking and live birth vs. no live birth (OR=5.89, 95% CI: 0.7, 50.1) (i.e. women who reported smoking at preconception were more likely to have a live birth). Conversely, ORs based on prospective reporting suggested a null association between smoking and live birth (OR=0.96, 95% CI: 0.4, 2.4).

Findings were similar when conducted using the expanded definitions for nonuse (data not shown). In this analysis we observed that 29% of women with no live birth in the study misclassified smoking status versus 16% of women with a live birth. We applied these levels of misclassification to a hypothetical population of 500 women with varying associations between smoking and live birth (OR=0.8, 1.0, and 1.5). In each case, we observed that misclassification had a strong affect on inferences. Specifically, for the scenario with a null association between smoking and live birth (OR=1.0), misclassification resulted in a significant underestimate (OR=0.55, 95% CI: 0.35, 0.87). When the true association was

protective (OR=0.8), misclassification resulted in an estimate biased away from the null (OR=0.51, 95% CI: 0.32, 0.80), whereas for the case where the true association was harmful (OR=1.5), misclassification resulted in a strong bias towards the null (OR=0.67, 95% CI: 0.43, 1.06).

Discussion

We observed that validity of maternal periconception behaviors was moderate at best and varied by type of behavior, though validity tended to be better for behaviors that may be more consistent (e.g., cigarette smoking) rather than inconsistent (e.g., alcohol consumption). Such misclassification may influence inferences regarding periconception behaviors and pregnancy outcomes, and future efforts should focus on prospective data collection. To our knowledge, this is the first validation study designed to empirically assess the validity of retrospectively reported data on women's behaviors during the periconception period that are purported to affect pregnancy outcomes. Such data are critical when relying on women's retrospective reporting of behavior during the periconception or early implantation stages of pregnancy, given increasing recognition of the importance of these sensitive windows for successful human development.^{30, 31}

While available periconception guidance underscores the importance of women engaging in healthy behaviors,^{32, 33} little evaluation of women's compliance has been undertaken.⁵ Our study is the first to assess women's recall of periconception behavior, which has implications for studies relying upon retrospectively reported behavioral data. Previous research has focused on validity of recall of pregnancy behaviors on birth and childhood outcomes,³⁴⁻⁴¹ and observed low⁴²⁻⁴⁴ to satisfactory or stable⁴⁵⁻⁴⁹ agreement for smoking and alcohol consumption. Similar to our study, smoking behaviors tended to show higher validity than alcohol consumption.⁵⁰⁻⁵² The use of prospective daily diaries to capture periconception behaviors is a unique aspect of our study, as other studies have relied on a single questionnaires during pregnancy^{53, 54} medical records^{55, 56} or interviews with midwives and family relatives^{57, 58} as the gold standard for assessing validity and reliability, although cotinine⁵⁹ biomarkers have also been utilized.

Our study offers two important findings: 1) validity is moderate at best; 2) validity varies by type of behavior. With the continual interest in the early life influences of later diseases and the intergenerational influences on health⁶⁰⁻⁶⁴ including gynecologic conditions collectively referred to as the ovarian dysgenesis hypothesis,¹⁴ it is imperative that investigators have some estimates about the validity of retrospectively reported behaviors during early sensitive windows and factors that may affect validity.⁶⁵ While validity was greatest for more regular type of behaviors such as cigarette smoking and consumption of caffeinated beverages, we were unable to identify any covariates that affected validity, with the exception of parity. Higher validity was observed for nulliparous rather than parous women. This finding may possibly reflect either the women's misreporting for the correct pregnancy, or her inability to remember correctly in the absence of adverse outcomes such as inability to become pregnant or carry a pregnancy to term. As discussed, our follow-up letters and questionnaires were tailored to each woman and her pregnancy outcome in the cohort study to ensure reporting for the correct pregnancy attempt. Given that agreement did not vary

consistently for all behaviors by pregnancy outcome, our findings do not support the second explanation noted above for the difference by parity. While considerable, reporting errors were bidirectional with women over- and under- reporting behaviors, suggesting that misclassification bias may be differential and generating bias in either direction. Thus, studies relying on retrospective reporting of this information are capable of either over- or under-estimating effect sizes.

This study has several important strengths including a gold standard based upon daily diaries with data captured well before pregnancy outcomes, coupled with a high participation rate for the follow-up survey conducted a decade later. Moreover, we found no systematic differences in participation status with regard to pregnancy outcome, behaviors, or other study covariates, suggesting that response bias would not account for our findings. Still, our findings require cautious interpretation given the limited size of our original cohort and our inability to query women more thoroughly about daily (rather than average) patterns of behaviors while trying to conceive. Additionally, given the small sample size we were unable to determine whether other demographic factors may be associated with validity. Other potential explanations for the relatively low validity observed in this study may be related to the length of follow-up, as retrospective reporting may have higher validity for shorter intervals such as during or shortly after pregnancy or perhaps several years following delivery.

Given the concerted effort to identify exposures during critical windows of human reproduction in relation to gynecologic and later onset adult diseases, many investigators rely upon retrospectively reported data from women obtained years, if not decades, later. Our findings suggest the need for considerable caution in interpreting such findings. Lastly, we cannot rule out the possibility that some behaviors associated with adverse pregnancy outcomes when relying upon retrospective reports may generate erroneous conclusions about etiology, given the observed bidirectional nature of reporting errors and the influence of misreporting on effect estimates of maternal smoking status and live birth. Despite our limited sample size, we observed that the levels of misclassification observed in this study could have a strong impact on inferences. Still, prospective cohort studies with periconception enrollment and longitudinal data capture are not always possible and support the need for continued development of methods to maximize the validity of behaviors or other exposures. Corroborating methodologically-oriented studies aimed at empirically estimating validity and reliability of self-reported data during this critical window of human develop remains a critical data gap in understanding the etiology of the early life influences of disease as well as helping clinicians evaluate the validity of recalled data on the part of their patients.

In conclusion, considerable variation in the validity of retrospectively reported periconception behaviors was observed, though validity tended to be better for behaviors that may be more consistent (e.g., cigarette smoking) rather than inconsistent (e.g., alcohol consumption); however, fish consumption did not follow this observation as its use was sporadic in this cohort. Such misclassification may generate bias leading to erroneous conclusions about the etiologic relation between recalled behaviors and pregnancy outcomes in comparison to prospectively collected data. Continued efforts to promote preconception

guidance may educate women about potential exposures while implicitly enhancing recall when queried later in pregnancy or at later points in life.

Acknowledgments

We would like to thank the Fertility and Fecundity Working Group at the *Eunice Kennedy Shriver* NICHD Division of Epidemiology, Statistics and Prevention Research for their guidance concerning this work.

Funding: This work was supported by the Agency for Toxic Substances and Disease Registry [H751 ATH 298338] and the Intramural Research Program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development at the National Institutes of Health.

List of Abbreviations and Acronyms

CI	Confidence interval
OR	Odds ratio
NICHD	<i>Eunice Kennedy Shriver</i> National Institute of Child Health and Human Development
TTP	Time-to-pregnancy

Reference List

1. Louis GM, Cooney MA, Lynch CD, Handal A. Periconception window: advising the pregnancy-planning couple. *Fertil.Steril.* 2008; 89:e119–e121. [PubMed: 18308052]
2. Johnson K, Posner SF, Biermann J, Cordero JF, Atrash HK, Parker CS, et al. Recommendations to improve preconception health and health care--United States. A report of the CDC/ATSDR Preconception Care Work Group and the Select Panel on Preconception Care. *MMWR Recomm.Rep.* 2006; 55:1–23. [PubMed: 16617292]
3. ACOG technical bulletin. Preconceptional care. Number 205--May 1995. American College of Obstetricians and Gynecologists. *Int.J.Gynaecol.Obstet.* 1995; 50:201–207. [PubMed: 7589761]
4. Thompson SJ, Torres ME, Stevenson RE, Dean JH, Best RG. Periconceptional multivitamin folic acid use, dietary folate, total folate and risk of neural tube defects in South Carolina. *Ann.Epidemiol.* 2003; 13:412–418. [PubMed: 12875798]
5. Inskip HM, Crozier SR, Godfrey KM, Borland SE, Cooper C, Robinson SM. Women's compliance with nutrition and lifestyle recommendations before pregnancy: general population cohort study. *BMJ.* 2009; 338:b481. [PubMed: 19213768]
6. Tough S, Tofflemire K, Clarke M, Newburn-Cook C. Do women change their drinking behaviors while trying to conceive? An opportunity for preconception counseling. *Clin.Med.Res.* 2006; 4:97–105. [PubMed: 16809401]
7. Kost K, Landry DJ, Darroch JE. The effects of pregnancy planning status on birth outcomes and infant care. *Fam.Plann.Perspect.* 1998; 30:223–230. [PubMed: 9782045]
8. Mohllajee AP, Curtis KM, Morrow B, Marchbanks PA. Pregnancy intention and its relationship to birth and maternal outcomes. *Obstet.Gynecol.* 2007; 109:678–686. [PubMed: 17329520]
9. Sable MR, Spencer JC, Stockbauer JW, Schramm WF, Howell V, Herman AA. Pregnancy wantedness and adverse pregnancy outcomes: differences by race and Medicaid status. *Fam.Plann.Perspect.* 1997; 29:76–81. [PubMed: 9099571]
10. Kaufmann RB, Morris L, Spitz AM. Comparison of two question sequences for assessing pregnancy intentions. *Am.J.Epidemiol.* 1997; 145:810–816. [PubMed: 9143211]
11. Peterson LS, Mosher WD. Options for measuring unintended pregnancy in cycle 6 of the National Survey of Family Growth. *Fam.Plann.Perspect.* 1999; 31:252–253. [PubMed: 10723655]

12. Skakkebaek NE, Rajpert-De ME, Main KM. Testicular dysgenesis syndrome: an increasingly common developmental disorder with environmental aspects. *Hum.Reprod.* 2001; 16:972–978. [PubMed: 11331648]
13. Smith GD. Assessing intrauterine influences on offspring health outcomes: can epidemiological studies yield robust findings? *Basic Clin.Pharmacol.Toxicol.* 2008; 102:245–256. [PubMed: 18226080]
14. Buck Louis, GM.; Cooney, MA. Effects of environmental contaminants on ovarian function and fertility. In: Gonzalez-Bulnes, A., editor. *Novel Concepts in Ovarian Endocrinology*. Kerala, India: Research Signpost; 2007. p. 249-268.
15. American College of Obstetricians and Gynecologists. *Good Health Before Pregnancy: Preconceptional Care*. 2007 Ref Type: Pamphlet.
16. Johnson K, Posner SF, Biermann J, Cordero JF, Atrash HK, Parker CS, et al. Recommendations to improve preconception health and health care--United States. A report of the CDC/ATSDR Preconception Care Work Group and the Select Panel on Preconception Care. *MMWR Recomm.Rep.* 2006; 55:1–23. [PubMed: 16617292]
17. Jacobson SW, Jacobson JL, Sokol RJ, Martier SS, Ager JW, Kaplan MG. Maternal recall of alcohol, cocaine, and marijuana use during pregnancy. *Neurotoxicol.Teratol.* 1991; 13:535–540. [PubMed: 1758408]
18. Yawn BP, Suman VJ, Jacobsen SJ. Maternal recall of distant pregnancy events. *J.Clin.Epidemiol.* 1998; 51:399–405. [PubMed: 9619967]
19. Post A, Gilljam H, Bremberg S, Galanti MR. Maternal smoking during pregnancy: a comparison between concurrent and retrospective self-reports. *Paediatr.Perinat.Epidemiol.* 2008; 22:155–161. [PubMed: 18298690]
20. Jacobson SW, Chiodo LM, Sokol RJ, Jacobson JL. Validity of maternal report of prenatal alcohol, cocaine, and smoking in relation to neurobehavioral outcome. *Pediatrics.* 2002; 109:815–825. [PubMed: 11986441]
21. Rice F, Lewis A, Harold G, van den BM, Boivin J, Hay DF, et al. Agreement between maternal report and antenatal records for a range of pre and peri-natal factors: the influence of maternal and child characteristics. *Early Hum.Dev.* 2007; 83:497–504. [PubMed: 17071023]
22. Jaspers M, de MG, Verhulst FC, Ormel J, Reijneveld SA. Limited validity of parental recall on pregnancy, birth, and early childhood at child age 10 years. *J.Clin.Epidemiol.* 2010; 63:185–191. [PubMed: 19682856]
23. Parazzini F, Davoli E, Rabaiotti M, Restelli S, Stramare L, Dindelli M, et al. Validity of self-reported smoking habits in pregnancy: a saliva cotinine analysis. *Acta Obstet.Gynecol.Scand.* 1996; 75:352–354. [PubMed: 8638455]
24. Czeizel AE, Petik D, Puho E. Smoking and alcohol drinking during pregnancy. The reliability of retrospective maternal self-reported information. *Cent.Eur.J.Public Health.* 2004; 12:179–183. [PubMed: 15666453]
25. Johnson K, Posner SF, Biermann J, Cordero JF, Atrash HK, Parker CS, et al. Recommendations to improve preconception health and health care--United States. A report of the CDC/ATSDR Preconception Care Work Group and the Select Panel on Preconception Care. *MMWR Recomm.Rep.* 2006; 55:1–23. [PubMed: 16617292]
26. Buck Louis GM, Dmochowski J, Lynch C, Kostyniak P, McGuinness BM, Vena JE. Polychlorinated biphenyl serum concentrations, lifestyle and time-to-pregnancy. *Hum.Reprod.* 2009; 24:451–458. [PubMed: 18940895]
27. Buck Louis GM, Dmochowski J, Lynch C, Kostyniak P, McGuinness BM, Vena JE. Polychlorinated biphenyl serum concentrations, lifestyle and time-to-pregnancy. *Hum.Reprod.* 2009; 24:451–458. [PubMed: 18940895]
28. Knaus H. Eine neue Methods zur Bestimmung des Ovulationstermines. *Zentralbl F Gynak.* 1929; 53:2193.
29. Ogino K. Ovulationstermin und Konzeptionstermin. *Zentralbl F Gynak.* 1930; 54:464–479.
30. Johnson K, Posner SF, Biermann J, Cordero JF, Atrash HK, Parker CS, et al. Recommendations to improve preconception health and health care--United States. A report of the CDC/ATSDR

- Preconception Care Work Group and the Select Panel on Preconception Care. *MMWR Recomm.Rep.* 2006; 55:1–23. [PubMed: 16617292]
31. Louis GM, Cooney MA, Lynch CD, Handal A. Periconception window: advising the pregnancy-planning couple. *Fertil.Steril.* 2008; 89:e119–e121. [PubMed: 18308052]
 32. ACOG technical bulletin. Preconceptional care. Number 205--May 1995. American College of Obstetricians and Gynecologists. *Int.J.Gynaecol.Obstet.* 1995; 50:201–207. [PubMed: 7589761]
 33. Johnson K, Posner SF, Biermann J, Cordero JF, Atrash HK, Parker CS, et al. Recommendations to improve preconception health and health care--United States. A report of the CDC/ATSDR Preconception Care Work Group and the Select Panel on Preconception Care. *MMWR Recomm.Rep.* 2006; 55:1–23. [PubMed: 16617292]
 34. Jacobson SW, Jacobson JL, Sokol RJ, Martier SS, Ager JW, Kaplan MG. Maternal recall of alcohol, cocaine, and marijuana use during pregnancy. *Neurotoxicol.Teratol.* 1991; 13:535–540. [PubMed: 1758408]
 35. Yawn BP, Suman VJ, Jacobsen SJ. Maternal recall of distant pregnancy events. *J.Clin.Epidemiol.* 1998; 51:399–405. [PubMed: 9619967]
 36. Post A, Gilljam H, Bremberg S, Galanti MR. Maternal smoking during pregnancy: a comparison between concurrent and retrospective self-reports. *Paediatr.Perinat.Epidemiol.* 2008; 22:155–161. [PubMed: 18298690]
 37. Jacobson SW, Chiodo LM, Sokol RJ, Jacobson JL. Validity of maternal report of prenatal alcohol, cocaine, and smoking in relation to neurobehavioral outcome. *Pediatrics.* 2002; 109:815–825. [PubMed: 11986441]
 38. Rice F, Lewis A, Harold G, van den BM, Boivin J, Hay DF, et al. Agreement between maternal report and antenatal records for a range of pre and peri-natal factors: the influence of maternal and child characteristics. *Early Hum.Dev.* 2007; 83:497–504. [PubMed: 17071023]
 39. Jaspers M, de MG, Verhulst FC, Ormel J, Reijneveld SA. Limited validity of parental recall on pregnancy, birth, and early childhood at child age 10 years. *J.Clin.Epidemiol.* 2010; 63:185–191. [PubMed: 19682856]
 40. Parazzini F, Davoli E, Rabaiotti M, Restelli S, Stramare L, Dindelli M, et al. Validity of self-reported smoking habits in pregnancy: a saliva cotinine analysis. *Acta Obstet.Gynecol.Scand.* 1996; 75:352–354. [PubMed: 8638455]
 41. Czeizel AE, Petik D, Puho E. Smoking and alcohol drinking during pregnancy. The reliability of retrospective maternal self-reported information. *Cent.Eur.J.Public Health.* 2004; 12:179–183. [PubMed: 15666453]
 42. Czeizel AE, Petik D, Puho E. Smoking and alcohol drinking during pregnancy. The reliability of retrospective maternal self-reported information. *Cent.Eur.J.Public Health.* 2004; 12:179–183. [PubMed: 15666453]
 43. Jacobson SW, Jacobson JL, Sokol RJ, Martier SS, Ager JW, Kaplan MG. Maternal recall of alcohol, cocaine, and marijuana use during pregnancy. *Neurotoxicol.Teratol.* 1991; 13:535–540. [PubMed: 1758408]
 44. Jaspers M, de MG, Verhulst FC, Ormel J, Reijneveld SA. Limited validity of parental recall on pregnancy, birth, and early childhood at child age 10 years. *J.Clin.Epidemiol.* 2010; 63:185–191. [PubMed: 19682856]
 45. Parazzini F, Davoli E, Rabaiotti M, Restelli S, Stramare L, Dindelli M, et al. Validity of self-reported smoking habits in pregnancy: a saliva cotinine analysis. *Acta Obstet.Gynecol.Scand.* 1996; 75:352–354. [PubMed: 8638455]
 46. Rice F, Lewis A, Harold G, van den BM, Boivin J, Hay DF, et al. Agreement between maternal report and antenatal records for a range of pre and peri-natal factors: the influence of maternal and child characteristics. *Early Hum.Dev.* 2007; 83:497–504. [PubMed: 17071023]
 47. Jacobson SW, Chiodo LM, Sokol RJ, Jacobson JL. Validity of maternal report of prenatal alcohol, cocaine, and smoking in relation to neurobehavioral outcome. *Pediatrics.* 2002; 109:815–825. [PubMed: 11986441]
 48. Yawn BP, Suman VJ, Jacobsen SJ. Maternal recall of distant pregnancy events. *J.Clin.Epidemiol.* 1998; 51:399–405. [PubMed: 9619967]

49. Post A, Gilljam H, Bremberg S, Galanti MR. Maternal smoking during pregnancy: a comparison between concurrent and retrospective self-reports. *Paediatr.Perinat.Epidemiol.* 2008; 22:155–161. [PubMed: 18298690]
50. Rice F, Lewis A, Harold G, van den BM, Boivin J, Hay DF, et al. Agreement between maternal report and antenatal records for a range of pre and peri-natal factors: the influence of maternal and child characteristics. *Early Hum.Dev.* 2007; 83:497–504. [PubMed: 17071023]
51. Jaspers M, de MG, Verhulst FC, Ormel J, Reijneveld SA. Limited validity of parental recall on pregnancy, birth, and early childhood at child age 10 years. *J.Clin.Epidemiol.* 2010; 63:185–191. [PubMed: 19682856]
52. Jacobson SW, Chiodo LM, Sokol RJ, Jacobson JL. Validity of maternal report of prenatal alcohol, cocaine, and smoking in relation to neurobehavioral outcome. *Pediatrics.* 2002; 109:815–825. [PubMed: 11986441]
53. Jacobson SW, Chiodo LM, Sokol RJ, Jacobson JL. Validity of maternal report of prenatal alcohol, cocaine, and smoking in relation to neurobehavioral outcome. *Pediatrics.* 2002; 109:815–825. [PubMed: 11986441]
54. Jaspers M, de MG, Verhulst FC, Ormel J, Reijneveld SA. Limited validity of parental recall on pregnancy, birth, and early childhood at child age 10 years. *J.Clin.Epidemiol.* 2010; 63:185–191. [PubMed: 19682856]
55. Rice F, Lewis A, Harold G, van den BM, Boivin J, Hay DF, et al. Agreement between maternal report and antenatal records for a range of pre and peri-natal factors: the influence of maternal and child characteristics. *Early Hum.Dev.* 2007; 83:497–504. [PubMed: 17071023]
56. Yawn BP, Suman VJ, Jacobsen SJ. Maternal recall of distant pregnancy events. *J.Clin.Epidemiol.* 1998; 51:399–405. [PubMed: 9619967]
57. Czeizel AE, Petik D, Puho E. Smoking and alcohol drinking during pregnancy. The reliability of retrospective maternal self-reported information. *Cent.Eur.J.Public Health.* 2004; 12:179–183. [PubMed: 15666453]
58. Post A, Gilljam H, Bremberg S, Galanti MR. Maternal smoking during pregnancy: a comparison between concurrent and retrospective self-reports. *Paediatr.Perinat.Epidemiol.* 2008; 22:155–161. [PubMed: 18298690]
59. Parazzini F, Davoli E, Rabaiotti M, Restelli S, Stramare L, Dindelli M, et al. Validity of self-reported smoking habits in pregnancy: a saliva cotinine analysis. *Acta Obstet.Gynecol.Scand.* 1996; 75:352–354. [PubMed: 8638455]
60. Dumesic DA, Schramm RD, Abbott DH. Early origins of polycystic ovary syndrome. *Reprod.Fertil.Dev.* 2005; 17:349–360. [PubMed: 15745643]
61. Smith GD. Intergenerational influences on health: how far back do we have to go? *Int.J.Epidemiol.* 2009; 38:617–618. [PubMed: 19504716]
62. Rinaudo PF, Lamb J. Fetal origins of perinatal morbidity and/or adult disease. *Semin.Reprod.Med.* 2008; 26:436–445. [PubMed: 18825611]
63. Davey SG, Leary S, Ness A, Lawlor DA. Challenges and novel approaches in the epidemiological study of early life influences on later disease. *Adv.Exp.Med.Biol.* 2009; 646:1–14. [PubMed: 19536658]
64. Gluckman PD, Hanson MA, Bateson P, Beedle AS, Law CM, Bhutta ZA, et al. Towards a new developmental synthesis: adaptive developmental plasticity and human disease. *Lancet.* 2009; 373:1654–1657. [PubMed: 19427960]
65. Smith GD. Assessing intrauterine influences on offspring health outcomes: can epidemiological studies yield robust findings? *Basic Clin.Pharmacol.Toxicol.* 2008; 102:245–256. [PubMed: 18226080]

Table 1

Validity of retrospectively reported behaviors in the periconception window – comparison of prospective* and retrospective data.

Prospectively Measured Behaviors*	Retrospectively Reported Behaviors			Based on Strict Definition of Nonuse		
	No n (%)	Yes n (%)	Kappa	No n (%)	Yes n (%)	% Agreement
Cigarette smoking						
No	59 (72.0)	0 (0)	0.43	59 (72.0)	0 (0)	81.8
Yes	15 (18.3)	8 (9.8)		15 (18.3)	8 (9.8)	
Fish consumption						
No	50 (63.3)	14 (17.7)	0.32	50 (63.3)	14 (17.7)	74.7
Yes	6 (7.6)	9 (11.4)		6 (7.6)	9 (11.4)	
Caffeine consumption						
No	2 (2.4)	3 (3.7)	0.21	2 (2.4)	3 (3.7)	85.4
Yes	9 (11.0)	68 (82.9)		9 (11.0)	68 (82.9)	
Vitamin usage						
No	0 (0)	0 (0)	--	0 (0)	0 (0)	74.2
Yes	17 (25.8)	49 (74.2)		17 (25.8)	49 (74.2)	
Alcohol consumption						
No	15 (18.3)	1 (1.2)	0.20	15 (18.3)	1 (1.2)	53.7
Yes	37 (45.1)	29 (35.4)		37 (45.1)	29 (35.4)	

* Prospectively measured behaviors are considered the gold standard.

Table II

Validity stratified by time to pregnancy and pregnancy outcome.

	TTP ≤ 6 cycles (n=52)			TTP > 6 cycles (n=10)		
	Kappa	Kappa [95% CI]	% Agreement	Kappa	Kappa [95% CI]	% Agreement
Cigarette smoking	0.57	[0.25, 0.89]	90	0.44	[0.01, 0.88]	70
Fish consumption	0.36	[0.08, 0.64]	76	0.55		78
Caffeine consumption	0.27	[-0.07, 0.61]	85	--*		90
Vitamin usage	--	--	70	--		78
Alcohol consumption	0.16	[0.01, 0.30]	48	--		60
			Live Birth (n=49)			Pregnancy Loss (n=13)
Cigarette smoking	0.59	[0.33, 0.85]	86	--		92
Fish consumption	0.55	[0.28, 0.83]	83	0.03	[-0.39, 0.44]	54
Caffeine consumption	0.33	[-0.07, 0.74]	88	--		77
Vitamin usage	--	--	73	--		67
Alcohol consumption	0.13	[-0.02, 0.28]	47	0.32	[-0.03, 0.66]	62
						No Conception (n=9)
						67
						67
						89
						100
						89

* -- not calculable because either row or column totals were 0 upon stratification.

Table III

Correlations for prospectively and retrospectively reported data restricting by abstinence.*

Lifestyle Behavior (#)	Total Sample			Excluding Abstainers		
	n	Correlation coefficient	P-value	n	Correlation coefficient	P-value
Daily cigarettes smoked	60	0.71	<0.0001	22	0.72	0.0002
Monthly Great Lakes fish meals	56	0.41	0.002	39	0.31	0.06
Daily caffeinated beverages	56	0.54	<0.0001	52	0.58	<0.0001
Weekly alcoholic drinks	58	0.39	0.003	53	0.36	0.008

* P-values were calculated for Spearman correlations coefficients. Abstainers comprised women who reported never having engaged in the behavior in the 12 months prior to enrollment into the prospective cohort.