Original Investigation

Women who remember, women who do not: A methodological study of maternal recall of smoking in pregnancy

Kate E. Pickett, Kristen Kasza, Gretchen Biesecker, Rosalind J. Wright, & Lauren S. Wakschlag

Abstract

Introduction: Retrospective recall of smoking during pregnancy is assumed to be substantially biased, but this has rarely been tested empirically.

Methods: We examined the validity of an interview-based retrospective recall more than a decade after pregnancy, in a cohort with repeated, multimethod characterization of pregnancy smoking ($N=245$). Retrospective smoking patterns were examined in relation to prospective reported and biological estimates of overall and trimester-specific smoking status and intensity. We also compared characteristics of women whose smoking status was misclassified by either prospective or retrospective measures with women whose status was congruent for non-smoking across timepoints.

Results: In general, sensitivity and specificity of recalled smoking were excellent relative to both prospective self-reported and cotinine-validated smoking status and trimester-specific intensity. However, measures were less congruent for amount smoked for women who recalled being heavy smokers. Further, retrospective measures captured some smokers not identified prospectively due to smoking that occurred prior to assessments. Women who would have been misclassified as nonsmokers based on either prospective or retrospective assessment differed significantly from congruently classified nonsmokers in a number of maternal, family, and neighborhood, but not child behavior, characteristics.

Discussion: When epidemiological studies of the impact of smoking in pregnancy use retrospective methods, misclassification may not be a significant problem if prenatal smoking is assessed in terms of the pattern across pregnancy. This type of

interview-based recall of pregnancy smoking may be relatively accurate, although optimal measurement should combine retrospective and prospective self-report and biological assays, as each provide unique information and sources of error.

Introduction

Epidemiological studies of the impact of maternal smoking in pregnancy on outcomes in offspring frequently rely on maternal retrospective self-report of smoking, sometimes soon after the pregnancy, sometimes many years later. It is often assumed that such retrospective reports are rife with recall error and bias. Yet, only a few studies have empirically examined this question in relation to prenatal smoking (Githens, Glass, Sloan, & Entman, 1993; Heath et al., 2003; Tomeo et al., 1999; Yawn, Suman, & Jacobsen, 1998). All report high levels of agreement between retrospective and prospective measures. However, limitations of measurement in these studies are problematic because none examined recall of smoking in a way that captures patterns of smoking across pregnancy. Among the few studies that have examined smoking, all used dichotomous measures of smoked versus not. Further, we are not aware of studies that have examined this question in relation to timing or intensity of smoking, the correspondence of recall to biological assays of maternal smoking, or factors that predict poor recall.

Measurement error that contributes to misclassification of smoking status may hamper detection of subtle or long-term effects of maternal smoking during pregnancy. For example, if a woman stopped smoking as soon as she learned she was pregnant, she may retrospectively report being a nonsmoker during pregnancy, truthfully from her own long-term perspective. She would be classified in most studies as a nonsmoker or quitter;

Kate E. Pickett, Ph.D., *Department of Health Sciences and Hull-York Medical School, University of York, UK*
Kristen Kasza, M.S., *Department of Health Studies, University of Chicago, IL*
Gretchen Biesecker, Ph.D., *Institute for Juvenile Research, Department of Psychiatry, University of Illinois, Chicago*
Rosalind J. Wright, M.D., M.P.H., *Channing Laboratory, Department of Medicine, Brigham Women's Hospital/Harvard Medical School, and Department of Environmental Health, Harvard School of Public Health, Boston, MA*

Lauren S. Wakschlag, Ph.D., *Institute for Juvenile Research, Department of Psychiatry, University of Illinois, Chicago*

Corresponding Author:

Kate E. Pickett, Ph.D., *Department of Health Sciences and Hull-York Medical School, University of York, Seabrook Rowntree Building, Area 3, Heslington, York YO10 5DD, UK. Telephone: +44-0-1904-321377; Fax: +44-0-1904-321388; E-mail: kp6@york.ac.uk*

doi: 10.1093/ntr/ntp117

Advance Access publication on July 28, 2009

Received February 22, 2009; accepted May 15, 2009

© The Author 2009. Published by Oxford University Press on behalf of the Society for Research on Nicotine and Tobacco.

All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org

yet her fetus would be at risk for any adverse outcomes associated with first-trimester exposure.

Smoking in pregnancy is a complex behavior that fluctuates over the course of pregnancy (Pickett, Rathouz, Kasza, Wakschlag, & Wright, 2005; Pickett, Wakschlag, Dai, & Leventhal, 2003). Although much emphasis has been placed on nondisclosure as a source of error, the extent to which mothers truthfully report whether or not they smoke is only one influence on the accuracy of measurement—others include the timing of measurement, comprehensiveness of assessment, and individual differences in smoking topography and metabolism. Retrospective interviews that query smoking across the course of pregnancy using methods designed to capture the manner in which mothers structure their behavior can substantially improve the reliability of reporting (Brigham et al., 2008). Similarly, biological assays are useful for obtaining a prospective direct measure of smoking but reflect approximately only the previous 24 hr.

In addition to the complex nature of smoking behavior and its impact on the validity of recall, it is likely that errors in maternal recall do not occur randomly, but little is known of the maternal and family characteristics that influence them. Studies have shown that recall of past events other than smoking may be influenced by maternal characteristics such as mood, personality, and mental health problems (Jaffee & Price, 2007). For example, maternal antisocial tendencies may be associated with nondisclosure or underreporting because lying is a characteristic of antisocial behavior, or alternatively with more accurate reporting because of a lack of concern with conforming to social mores. Maternal depression has also been associated with systematic reporting errors. Maternal socioeconomic status may also influence validity of reporting due to varying social norms around smoking in different socioeconomic and social class contexts (Pickett, Wakschlag, Rathouz, Leventhal, & Abrams, 2002). Having a child with a mental health problem may also introduce bias in recall (Jaffee & Price, 2007).

The direction and magnitude of bias caused by the misclassification of smoking during pregnancy will depend on the magnitude of the true effect of exposure, as well as the degree of misclassification. Under most plausible scenarios, however, misclassification is likely to lead to an underestimate of the true risk. Previously reported simulations of the effects of misclassifying smoking during pregnancy in epidemiological studies found that the underestimation of the relative risk for smoking on hypothesized adverse outcomes ranged from <10% to 55% under varying misclassification scenarios (Pickett et al., 2003).

In this study, we address gaps in research on the validity and reliability of maternal recall of smoking within a unique cohort with repeated prospective measurements of smoking that include both self-report and biological assays, along with retrospective recall of pregnancy smoking patterns and rich data on maternal and family characteristics and social context. We link these to address the following research questions:

1. What is the congruence between maternal retrospective report of prenatal smoking and prospective measurement? Does this vary by measurement method, timing, and intensity of smoking?
2. Are maternal, family, neighborhood, and child behavior characteristics related to the misclassification of smoking by prospective and retrospective reports?

Methods

Data and study sample

For these analyses, we used data collected during a follow-up study of a prospective cohort of pregnant women and infants who were enrolled in 1986–1992. The study was originally established as the Maternal–Infant Smoking Study of East Boston to compare the relative effects of in utero and early-life exposure to cigarette smoke on infant lung function (Hanrahan et al., 1992; Tager, Ngo, & Hanrahan, 1995). The follow-up was conducted as the East Boston Family Study (EBFS) as the offspring entered adolescence, with a primary aim of investigating the long-term impact of smoking in pregnancy on offspring behavior (Wakschlag et al., 2009). Follow-up data presented here are drawn from Wave 1 of EBFS. In this paper, pregnancies are the unit of analysis, and eligibility is restricted to the 245 pregnancies where the biological mother was interviewed and smoking information was available in both the original study and the follow-up study. (Detailed descriptions of the sampling and eligibility criteria are available in Supplementary Materials.)

Prospective measurement of smoking and cotinine levels during pregnancy

Women were recruited at the first prenatal visit; the timing of this varied from 10 to 27 weeks gestation, with 21% recruited in the first trimester. They were asked at baseline about their past and current smoking status. At each subsequent prenatal visit (modal number of visits = 7, range 1–12), women reported current smoking habits, including the number of cigarettes being smoked per day. A urine sample was collected and analyzed for measurement of cotinine by radioimmunoassay at each of these visits. Urine cotinine values were corrected for urine concentration and expressed as nanogram per milliliter of urinary creatinine. We have reported previously on the complexity of concordance between prospective self-report and cotinine measures in this sample (Pickett et al., 2005). In brief, the sensitivity and specificity of prospective self-report compared with cotinine-indicated smoking were 88.4% and 99.0%, respectively.

Retrospective measurement of smoking

In the follow-up study, women were asked to recall their smoking behavior during pregnancy. The mean duration of recall was 14.5 years ($SD = 1.7$ years); study offspring were aged from 11 to 18 years. Women were asked how many cigarettes per day they smoked before they knew they were pregnant, in the first trimester after they knew they were pregnant, and in the second and third trimesters. They were also asked about current smoking habits.

Maternal characteristics

Measures of maternal sociodemographic characteristics at follow-up included age, employment status, receipt of public benefit, and home ownership. Mothers' health was measured as self-rated general health, self-rated emotional health, history of mental health service use, and history of psychiatric hospitalizations. Maternal problem behaviors included: substance misuse, measured as a score >6 on the Drug Abuse Screening Test (Skinner, 1982); problematic alcohol use, assessed as a score >4 on a modified 51-item version of the Michigan Alcoholism Screening Test (Selzer, 1971); and antisocial behavior, assessed

with a score >24 on the Antisocial Behavior Checklist (Zucker, Noll, Ham, Fitzgerald, & Sullivan, 1994). Personality traits related to problem behavior were measured using the Zuckerman–Kuhlman Personality Questionnaire, which gives mean scores for impulsivity/sensation seeking and aggression/hostility (Zuckerman, 2002). Measures of maternal stress and support included the mean score on the 16-item Parental Well-Being Index (Dunst & Trivette, 1986) and the 19-item Medical Outcomes Study Social Support Survey—higher scores on each are positive (Sherbourne & Stewart, 1991). We also measured stress in the past month using mean score on the four-item Perceived Stress Scale (Cohen, Kamarck, & Mermelstein, 1983) and lifetime stress with the mean score on a modified five-item Evaluation of Lifetime Stressors Questionnaire—higher scores on these indicate more stress (Krinsley, 1996).

Family characteristics

Marital status was coded dichotomously. The adequacy of material resources within the family was measured by mean score on the 30-item Family Resources Scale—high scores are positive (Dunst & Leet, 1985). Family instability was measured as the mean score of a 9-item scale, based on the work of Forman and Davies (2003), which measured changes in residence, primary caregiver, transitions in romantic relationships of the primary caregiver, job or income loss, and death or serious illness of a family member. High scores indicate more instability.

Neighborhood characteristics

Perceptions of neighborhood violence were measured with three items on a 5-point scale, developed specifically for the EBFS study; high scores indicate the mother experiencing more neighborhood violence.

Child behavior measures

We also examine youth conduct problems because of the focus of current research on causal modeling in relation to the smoking: Conduct problems link. Antisocial behavior was measured via the Antisocial Behavior Checklist, a 60-item inventory completed by the adolescent offspring about themselves (Zucker et al., 1994). Conduct disorder symptoms were measured by asking the parent about the adolescent offspring's behavior, using the Diagnostic Interview Schedule for Children (Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000).

Statistical analysis

Recall of smoking status. We first assessed the accuracy and reliability of women's long-term retrospective self-report of *smoking status* in pregnancy (i.e., smoked or not) by examining the sensitivity, specificity, and kappa statistics for retrospective report of any smoking in pregnancy in comparison with two prospective measures—ever having reported smoking during pregnancy and ever having a cotinine value indicating active smoking during pregnancy. Next, we reclassified 36 women who retrospectively reported smoking only before they knew they were pregnant and who (if they had truly not smoked once they learned of the pregnancy) could not have been classified as smokers prospectively at prenatal visits and reexamined sensitivity, specificity, and kappa statistics in comparison to prospective self-report of any smoking and positive cotinine measures.

Recall of timing of smoking. To examine whether or not recall is affected by the timing of smoking in pregnancy, we next assessed the reliability of maternal recall of smoking status during each trimester based on retrospective self-report versus prospective self-report and prospective cotinine measures.

Recall of smoking intensity. We examined recall of intensity of smoking using the Bland and Altman (1986) technique to assess the degree of agreement in number of cigarettes smoked between prospective and retrospective report among the 87 women for whom there were both retrospective and prospective evidences of smoking. We also examined whether there was a positive linear relationship between the two reports by examining the correlation between number of cigarettes smoked per day among the same 87 women; we looked at the correlations and cross-tabulations between the amount smoked averaged across pregnancy and within each trimester.

Predictors of misclassification by prospective and retrospective reports. Finally, we sought to understand how maternal, family, neighborhood, and child behavioral characteristics were related to the probability that a mother's smoking would be misclassified in an epidemiological study that used only prospective or retrospective measures. Women who recalled smoking in pregnancy but whose smoking was never captured prospectively by either self-report or cotinine were classified as retrospective-only smokers. In retrospective studies, these women would be classified as smokers, but in prospective studies as nonsmokers. We classified women who were captured as smoking by prospective measures but who did not recall smoking retrospectively as prospective-only smokers. In prospective studies, these women would be classified as smokers, but in retrospective studies as nonsmokers. As in either study design, incongruent smokers would be misclassified as nonsmokers, so we examine their characteristics with reference to congruent nonsmokers. Due to small numbers of women with incongruent measures, we interpret p values < .10 as statistically significant.

Results

Overall, women's long-term retrospective recall of smoking in pregnancy was accurate and reliable, in comparison with both prospective self-report and prospective biological assessment of smoking status (Table 1). In analyses examining recall of smoking status, among women who were positive for ever smoking in pregnancy by cotinine, 95.6% recalled having smoked, while among those who prospectively reported smoking, recall sensitivity was slightly higher, at 98.1%. Among women who were never identified as smokers by a positive cotinine prospectively, 19.7% recalled having smoked; specificity was slightly lower for self-report—among those who prospectively reported no smoking, 22.7% reported retrospectively that they had smoked during that pregnancy. This low specificity is improved (although sensitivity declined) when comparing recall of smoking in pregnancy after women knew they were pregnant with prospective measures. Kappa statistics showed excellent agreement for all assessments of the reliability of any smoking.

Table 2 shows the prevalence of recalled smoking for different time periods of pregnancy versus prospective self-report and cotinine measures for different time periods. Three aspects of

Table 1. Accuracy and reliability of maternal recall of smoking in pregnancy

	Ever positive cotinine (prospective)		
	Yes	No	Total
Recalled any smoking			
Yes	108	26	134
No	5	106	111
Total	113	132	245
	Sensitivity = 95.6%		Specificity = 80.3%
	Ever positive self-report (prospective)		$\kappa = .76$
Recalled any smoking	Yes	No	Total
Yes	102	32	134
No	2	109	111
Total	104	141	245
	Sensitivity = 98.1%		Specificity = 77.3%
	Ever positive cotinine (prospective)		$\kappa = .75$
Recalled any smoking after learning of pregnancy	Yes	No	Total
Yes	89	9	98
No	24	123	147
Total	113	132	245
	Sensitivity = 78.8%		Specificity = 93.2%
	Ever positive self-report (prospective)		$\kappa = .73$
Recalled any smoking after learning of pregnancy	Yes	No	Total
Yes	88	10	98
No	16	131	147
Total	104	141	245
	Sensitivity = 84.6%		Specificity = 92.9%
			$\kappa = .78$

these distributions are notable. First, 12% of women recalled smoking at some time during pregnancy but were never identified as smokers prospectively, either by self-report or by cotinine (retrospective-only smokers). A large majority of these women recalled smoking only before knowing they were pregnant or only in the first trimester. As prospective measures were made at prenatal care visits and relatively few visits were first-trimester visits, there were no prospective possibilities for picking up women who quit as soon as they learned they were pregnant and few opportunities for detecting women who smoked only in

early pregnancy. For this group of women, retrospective report of smoking may be capturing true smoking which no practical prospective measurement could capture. Second, 9% of women recalled either never having smoked or only before they knew they were pregnant but had been prospectively identified as smokers either by self-report or by cotinine at one or more prenatal visits (prospective-only smokers). For this group of women, retrospective report is clearly inferior to prospective measurement, particularly prospective measurement by cotinine. Also notable is that the congruence between retrospective

Table 2. Prevalence of smoking by time period, by recall vs. prospective measurement

Retrospective recall of smoking during pregnancy	N ^a	Prospective self-report of any smoking N ^a					Prospective cotinine positive for smoking N ^a				
		1	2	3	4	5	1	2	3	4	5
1. None	111	109	N/A	0	1	1	106	N/A	1	3	2
2. Smoked but not after learning of pregnancy	36	22	N/A	1	14	11	19	N/A	1	17	16
3. Smoked in first trimester	96	10	N/A	16	85	82	8	N/A	13	86	80
4. Smoked in second trimester	82	4	N/A	16	77	75	3	N/A	11	77	73
5. Smoked in third trimester	84	4	N/A	15	79	77	4	N/A	12	78	75

Note. Light gray shaded boxes denote recalled smoking but not identified as smoker prospectively (retrospective-only smoker), total number of women = 29 (12%). Dark gray shaded boxes denote recalled not smoking but identified as smoker prospectively (prospective-only smoker), total number of women = 22 (9%).

^aNumbers in table refer to number of reports, not number of women, that is, measurement is for four separate time periods for each woman. The numbers in each column reflect the number of smokers, captured with prospective measures. The numbers in rows reflect the number of smokers, captured with retrospective measures. Concordant measures are in the diagonal cells.

and prospective measurement for women who do recall smoking is much lower for the first trimester of pregnancy than for the second and third.

Table 3 compares average cigarette consumption per day for the subset of 87 congruent smokers for whom there was a retrospective report of smoking during pregnancy and prospective evidence of smoking by self-report. Retrospectively, 25 women (28.7%) reported very heavy smoking (more than a pack per day), whereas prospectively, only 15 (17.2%) were classified as very heavy smokers. Recall of light and moderate smoking agrees better with prospective categorization of smoking intensity than recall of heavy smoking. More than 60% of women who recall light or moderate smoking would have been similarly classified prospectively, and only 4 of these 62 women seem to be underreporting heavy smoking. In contrast, less than half the women who recall heavy smoking would have been prospectively classified as heavy smokers; almost half of them (56%) would have been classified as smoking less than, instead of more than, a pack per day. The mean difference between prospective and retrospective reported cigarettes per day was -0.523 , and the 95% *CI* for agreement was -2.14 to 1.10 cigarettes/day (Supplementary Figure 2).

The correlation between recalled average cigarettes per day versus prospective average number of cigarettes per day across pregnancy was moderate ($r = .58, p < .001$). Correlations of recalled versus prospective smoking intensity were somewhat stronger for the second ($r = .53, p < .001$) and third ($r = .57, p < .001$) trimester than for the first ($r = .41, p = .11$).

Of 23 maternal, family, neighborhood, or child behavior characteristics that we examined, 5 were significantly different for both prospective-only and retrospective-only smokers in relation to nonsmokers, and a further 4 were significantly different between retrospective-only smokers and nonsmokers (Table 4). Compared with nonsmokers, both prospective-only and retrospective-only smokers were significantly more likely to be receiving public benefits (odds ratio [*OR*] = 4.6, 95% *CI* = 1.7–12.1, $p < .01$ and *OR* = 3.4, 95% *CI* = 1.4–8.4, $p < .01$, respectively) and less likely to own their home (*OR* = 0.4, 95% *CI* = 0.1–0.9, $p = .04$ and *OR* = 0.1, 95% *CI* = 0.05–0.4, $p < .01$, respectively). They were more likely to have low self-rated health (*OR* = 4.5, 95% *CI* = 1.4–14.8, $p = .01$ and *OR* = 3.9, 95% *CI* = 1.3–11.8, $p = .02$, respectively), less likely to be married (*OR* = 0.4, 95% *CI* = 0.2–1.0, $p = .06$ and *OR* = 0.5, 95% *CI* =

0.2–1.0, $p = .06$, respectively), and scored their neighborhood higher on the violence scale (7.6 vs. 6.2, $p = .03$ and 7.3 vs. 6.2, $p = .07$, respectively).

Additionally, retrospective-only smokers were more likely than nonsmokers to have low emotional health (*OR* = 2.7, 95% *CI* = 1.0–7.3, $p = .05$) and to score lower on measures of parental well-being (42.9 vs. 46.7, $p = .08$), perceived stress (10.7 vs. 8.8, $p < .01$), and lifetime stressors (16.5 vs. 14.7, $p = .04$).

Neither group of misclassified smokers were significantly more likely to have children with Conduct Disorder symptoms or to have higher antisocial behavior scores than nonsmokers.

Discussion

Findings here indicate that women's ability to recall their smoking behavior in pregnancy more than a decade after the event is generally both accurate and reliable, particularly for the second and third trimester of pregnancy. The relative accuracy of retrospective reporting here is particularly striking in light of the detailed exposure measurement on which the prospective assessments were based.

Our results indicate that retrospective and prospective measures are less congruent in two instances: characterization of smoking during the first trimester and classification of women as very heavy smokers (pack per day or more). In both these instances, it was the *retrospective* assessment that was more informative than the prospective measures. The fact that more first-trimester smoking was identified retrospectively highlights the fact that timing of prospective assessments (e.g., they typically do not begin until well into the first trimester) may miss a key period of smoking for the substantial number of women who quit spontaneously upon learning they are pregnant and suggests that, even in prospective studies, queries about smoking patterns prior to pregnancy knowledge will be useful (i.e., "retrospective reports" of the few months prior). The fact that higher numbers of heavier smokers were identified retrospectively highlights the fact that underreporting may be substantial among pregnancy smokers. Based on calibration in relation to cotinine levels, a substantial number of smokers in this sample required adjustment in smoking intensity due to underreporting (Wakschlag et al., 2009). This pressure to underreport may be relaxed retrospectively.

Table 3. Recalled cigarette consumption per day in comparison with prospective reports of cigarettes per day, among women who smoked^a

Retrospective recall of average number of cigarettes per day during pregnancy	N	Prospective report of average number of cigarettes per day during pregnancy		
		Light (<1–9)	Moderate (10–19)	Heavy (20+)
Light (<1–9)	28	17 (60.7%)	11 (39.3%)	0 (0.0%)
Moderate (10–19)	34	9 (26.5)	21 (61.8)	4 (11.8)
Heavy (20+)	25	2 (8.0)	12 (48.0)	11 (44.0)
N		28	44	15

Note. ^aSubset of congruent smokers who recalled smoking at some point during pregnancy for whom there was also prospective evidence of smoking.

Table 4. Characteristics of congruent nonsmokers and congruent smokers and noncongruent prospective- or retrospective-only smokers

Characteristics	Congruent nonsmokers, <i>n</i> = 107	Prospective-only, <i>n</i> = 22	Retrospective-only, <i>n</i> = 29	Congruent smokers, <i>n</i> = 87
Maternal sociodemographic				
Mean age (<i>SD</i>)	42.2 (5.9)	41.4 (5.2)	42.8 (5.1)	40.6 (5.3)
Employed (%)	74.8	59.1	69.0	60.9
Receiving public benefits (%)	17.9	50.0 ^a	42.9 ^b	32.6
Own home (%)	57.0	36.4 ^a	17.2 ^b	39.1
Maternal health and well-being				
Low self-rated general health (%)	7.6	27.3 ^a	24.1 ^b	19.8
Low self-rated emotional health (%)	12.4	13.6	27.6 ^b	44.2
History of mental health service use (%)	47.6	45.4	51.7	50.0
History of psychiatric hospitalizations (%)	12.4	4.5	24.1	13.9
Substance use (%)	11.5	19.0	13.8	28.7
Problematic drinking (%)	23.3	35.0	25.0	40.7
Antisocial behavior (%)	8.6	14.3	17.2	14.9
Mean impulsive-sensation seeking score (<i>SD</i>)	4.8 (3.9)	5.1 (3.5)	6.0 (4.4)	5.4 (4.0)
Mean aggression-hostility score (<i>SD</i>)	4.5 (3.1)	4.9 (3.3)	5.2 (3.6)	6.1 (3.3)
Mean parental well-being score (<i>SD</i>)	46.7 (9.3)	48.4 (8.1)	42.9 (10.4) ^b	41.9 (11.1)
Mean social support score (<i>SD</i>)	4.2 (0.7)	4.3 (0.7)	4.3 (0.7)	4.1 (0.8)
Mean perceived stress score (<i>SD</i>)	8.8 (3.1)	9.2 (3.3)	10.7 (2.9) ^b	10.8 (3.4)
Mean lifetime stressors score (<i>SD</i>)	14.7 (3.7)	15.0 (3.6)	16.5 (3.7) ^b	15.6 (4.2)
Family characteristics				
Married (%)	67.3	45.4 ^a	48.3 ^b	65.5
Mean family resources score (<i>SD</i>)	4.3 (0.6)	4.2 (0.4)	4.1 (0.5)	4.0 (0.6)
Mean family instability score (<i>SD</i>)	0.4 (0.4)	0.5 (0.3)	0.5 (0.4)	0.5 (0.4)
Neighborhood quality				
Mean neighborhood violence score (<i>SD</i>)	6.2 (2.6)	7.6 (2.6) ^a	7.3 (2.9) ^b	6.8 (3.0)
Child behavior problems				
Mean conduct disorder symptoms (<i>SD</i>)	1.7 (2.7)	1.7 (2.9)	2.0 (1.9)	3.0 (3.6)
Mean antisocial behavior score (<i>SD</i>)	6.6 (7.3)	7.3 (8.1)	7.4 (6.6)	9.8 (10.5)

Note. ^aDifference between prospective-only vs. congruent nonsmokers statistically significant at *p* value < .10.

^bDifference between retrospective-only vs. congruent nonsmokers statistically significant at *p* value < .10.

Previously, researchers have shown that maternal recall of smoking in pregnancy agrees well with smoking being recorded in the medical record, over recall periods of 4–6 (Githens et al., 1993), 10–15 (Yawn et al., 1998), and 30+ years (Tomeo et al., 1999). Medical records may consist only of self-report of smoking, rather than any biomarker confirmation, and women who are more or less willing to disclose smoking prospectively in clinical settings may have similar tendencies toward disclosing smoking when asked to recall for a research study. However, using repeated prospective self-report and biological measures of smoking in pregnancy, we also find that most women are accurate and reliable reporters at an average of 14.5 years after the pregnancy.

Our prospective sample was recruited between 1986 and 1992. In the United States in 1990, the overall prevalence of smoking in pregnancy was 21%, but this differed greatly by region, ethnicity, and socioeconomic status. The population base for our study was chosen on the basis of high rates of smoking among White, working-class women (in our sample, 46% of women are classed as smokers by cotinine and 42% by prospective self-report). It is possible that, among this group, willingness to disclose smoking, either prospectively or retrospectively,

may differ from that among women of higher social class or from groups with different social norms around smoking in pregnancy. However, this seems unlikely to affect the patterns of concordance which were of primary interest in this study.

Our results confirm what Heath et al. (2003) reported using an innovative design: Self-report may be the best way to capture smoking in early pregnancy. In their study, Heath et al. asked women with twin sisters to recall smoking behavior during a particular pregnancy and then asked the twin sister to independently recall the women's smoking behavior. In general, the sisters' reports were in agreement, except for women who smoked early in pregnancy and then quit—the twin sisters under-reported this behavior.

It is relevant, however, that women who smoke during pregnancy are different, across multiple dimensions, from women who quit and women who never smoke, before or during pregnancy. Compared with women who smoke throughout pregnancy, women who do not smoke or who manage to quit smoking are more likely to be older, employed, and better educated (Agrawal et al., 2008; Graham & Der, 1999; Hanna, Faden, & Dufour, 1994; Martin et al., 2008). They are less likely to be

depressed (Hanna et al., 1994; Pritchard, 1994) or highly stressed (Paarlberg et al., 1999), to be a single or cohabiting parent (Kiernan & Pickett, 2006; Thue, Schei, & Jacobsen, 1995), to have a smoking partner (Appleton & Pharoah, 1998; Wakefield, Gillies, Graham, Madeley, & Symonds, 1993) or an abusive partner (McFarlane, Parker, & Soeken, 1996), to have low social support (Dejin-Karlsson et al., 1996), or to live in a working-class neighborhood (Pickett et al., 2002; Sellstrom, Arnoldsson, Bremberg, & Hjern, 2008). Three studies within the Family Health and Development Project have shown that women who quit and women who continue to smoke are systematically different across multiple domains of psychosocial problems, including interpersonal problems, adaptive functioning, and other health-risk behaviors (Wakschlag et al., 2003), as well as conduct problems (Kodl & Wakschlag, 2004), and problematic psychosocial context also had incremental utility for predicting smoking intensity (Weaver, Campbell, Mermelstein, & Wakschlag, 2008). These differences have been confirmed within the population-based U.K. Millennium Cohort Study (Pickett, Wilkinson, & Wakschlag, 2009). We are unaware of previous studies that have examined such characteristics in relation to recall of smoking in pregnancy.

Reassuringly, we found that on the majority of measures of maternal, family, or neighborhood context, women whose smoking status would be misclassified in prospective or retrospective studies did not differ significantly from the nonsmokers they would be grouped with. However, misclassified smokers did differ in some aspects, including sociodemographic characteristics, maternal health, and family and neighborhood context. In epidemiological studies of the impact of smoking in pregnancy on behavioral problems in offspring, misclassification may not lead to a biased estimate of the effect of smoking, as misclassified smokers had children with behavioral characteristics more similar to nonsmokers than to correctly classified smokers. However, we note that any residual bias would lead to slight underestimates of any effect; studies with significant numbers of misclassified smokers are likely to produce conservative estimates. Although our study had limited power to detect such differences, comparison of actual differences, as well as *p* values, supports this interpretation.

One study has reported agreement between maternal recall of smoking in pregnancy and the medical record in relation to child behavior problems (Rice et al., 2007). There was greater incongruence between reports among more educated women, but no evidence that child behavior problems affected congruence. Our findings confirm this with the added reassurance of prospective biological measures of smoking. In fact, our findings imply that retrospective measures may misclassify fewer smokers than prospective studies, which miss early quitters. One caveat, however, is that retrospective reports in the present study were interview based and queried trimester-by-trimester behavior (rather than being a single questionnaire item on “typical day” smoking). That is, all retrospective reports are not alike, and future examinations of this question should systematically test the optimal way for retrospective reports to be structured to elicit accurate recall. In addition, it should be noted that women in the present study had been participating in a “smoking study” for nearly two decades, including having provided detailed reports and biological samples repeatedly. Thus, they may have been unusually sensitized to remembering their previous smoking behavior.

In a study of the effects of misclassification of smoking during pregnancy, England et al. found that 21.6% of self-reported quitters had cotinine measures that were positive for smoking, and this was most common among women who reported quitting before pregnancy. Misclassification had modest effects on estimates of the impact of smoking. Nevertheless, this study shows that prospective biological assays of smoking are informative beyond prospective self-report, and we have shown in this paper that retrospective report is informative above and beyond both prospective self-report and biological assays. This suggests that optimal measurement of smoking in pregnancy will combine all three. Repeated prospective self-report and cotinine measures, in combination with detailed timeline followback measures of smoking before pregnancy, after conception but before pregnancy is known, in the early weeks of pregnancy, and for periods between prospective measures, would provide the most complete coverage of women’s smoking behavior over the course of pregnancy. Dukic, Niessner, Benowitz, Hans, and Wakschlag (2007) have proposed methods for combining prospective measures of self-report and cotinine—combining these with timeline followback measures may be the best way to characterize patterns of smoking in pregnancy. Although retrospective reporting has unique sources of error (such as recall bias), the novelty of the present findings is the implication that retrospective recall should not be treated as a flawed, last resort. Rather, these findings suggest that retrospective recall adds a unique source of information, perhaps due to reductions in the social sanctions that contribute to prospective nondisclosure and the opportunity it presents to report fully on the course of smoking across the pregnancy. Thus, we conclude that the optimal method for reducing risk of misclassification of pregnancy smoking status would combine prospective self-reported and biological measures as well as retrospective assessments.

Funding

The EBFS was supported by a grant from the National Institute on Drug Abuse to Dr. Wakschlag (R01DA15223). Dr. Wakschlag was also supported by the Walden and Jean Young Shaw and Children’s Brain Research Foundations. Dr. Pickett is supported by a U.K. National Institute of Health Research Career Scientist Award.

Declaration of Interests

None declared.

Supplementary material

Supplementary Figures 1–4 can be found at Nicotine and Tobacco Research online (<http://www.ntr.oxfordjournals.org/>).

References

- Agrawal, A., Knopik, V. S., Pergadia, M. L., Waldron, M., Bucholz, K. K., Martin, N. G., et al. (2008). Correlates of cigarette smoking during pregnancy and its genetic and environmental overlap with nicotine dependence. *Nicotine & Tobacco Research*, 10, 567–578.

- Appleton, P. L., & Pharoah, P. O. D. (1998). Partner smoking behavior change is associated with women's smoking reduction and cessation during pregnancy. *British Journal of Health Psychology*, 3, 361–374.
- Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1, 307–310.
- Brigham, J., Lessov-Schlaggar, C. N., Javitz, H. S., McElroy, M., Krasnow, R., & Swan, G. E. (2008). Reliability of adult retrospective recall of lifetime tobacco use. *Nicotine & Tobacco Research*, 10, 287–299.
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24, 385–396.
- Dejin-Karlsson, E., Hanson, B. S., Ostergren, P. O., Ranstam, J., Isacson, S. O., & Sjoberg, N. O. (1996). Psychosocial resources and persistent smoking in early pregnancy—a population study of women in their first pregnancy in Sweden. *Journal of Epidemiology and Community Health*, 50, 33–39.
- Dukic, V. M., Niessner, M., Benowitz, N., Hans, S., & Wakschlag, L. (2007). Modeling the relationship of cotinine and self-reported measures of maternal smoking during pregnancy: A deterministic approach. *Nicotine & Tobacco Research*, 9, 453–465.
- Dunst, C. J., & Leet, H. E. (1985). *Family Resources Scale: Reliability and validity*. Ashville, NC: Winterberry Press.
- Dunst, C. J., & Trivette, C. M. (1986). Looking beyond the parent-child dyad for the determinants of maternal styles of interaction. *Infant Mental Health Journal*, 7, 69–80.
- England, L. J., Grauman, A., Qian, C., Wilkins, D. G., Schisterman, E. F., Yu, K. F., & Levine, R. J. (2007). Misclassification of maternal smoking status and its effects on an epidemiologic study of pregnancy outcomes. *Nicotine and Tobacco Research*, 9, 1005–1013.
- Forman, E. M., & Davies, P. T. (2003). Family instability and young adolescent maladjustment: The mediating effects of parenting quality and adolescent appraisals of family security. *Journal of Clinical Child & Adolescent Psychology*, 32, 94–105.
- Githens, P. B., Glass, C. A., Sloan, F. A., & Entman, S. S. (1993). Maternal recall and medical records: An examination of events during pregnancy, childbirth, and early infancy. *Birth*, 20, 136–141.
- Graham, H., & Der, G. (1999). Patterns and predictors of smoking cessation among British women. *Health Promotion International*, 14, 231–239.
- Hanna, E. Z., Faden, V. B., & Dufour, M. C. (1994). The motivational correlates of drinking, smoking, and illicit drug use during pregnancy. *Journal of Substance Abuse*, 6, 155–167.
- Hanrahan, J. P., Tager, I. B., Segal, M. R., Tosteson, T. D., Castile, R. G., Van Vunakis, H., et al. (1992). The effect of maternal smoking during pregnancy on early infant lung function. *American Review of Respiratory Disease*, 145, 1129–1135.
- Heath, A. C., Knopik, V. S., Madden, P. A., Neuman, R. J., Lynskey, M. J., Slutske, W. S., et al. (2003). Accuracy of mothers' retrospective reports of smoking during pregnancy: Comparison with twin sister informant ratings. *Twin Research*, 6, 297–301.
- Jaffee, S. R., & Price, T. S. (2007). Gene-environment correlations: A review of the evidence and implications for prevention of mental illness. *Molecular Psychiatry*, 12, 432–442.
- Kiernan, K., & Pickett, K. E. (2006). Marital status disparities in maternal smoking during pregnancy, breastfeeding and maternal depression. *Social Science & Medicine*, 63, 335–346.
- Kodl, M. M., & Wakschlag, L. S. (2004). Does a childhood history of externalizing problems predict smoking during pregnancy? *Addictive Behaviors*, 29, 273–279.
- Krinsley, K. E. (1996). Psychometric review of the Evaluation of Lifetime Stressors (ELS) questionnaire and interview. In B. H. Stamm (Ed.), *Measurement of stress, trauma and adaptation* (pp. 160–162). Lutherville, MD: Sidran Press.
- Martin, L. T., McNamara, M., Milot, A., Bloch, M., Hair, E. C., & Halle, T. (2008). Correlates of smoking before, during, and after pregnancy. *American Journal of Health Behavior*, 32, 272–282.
- McFarlane, J., Parker, B., & Soeken, K. (1996). Physical abuse, smoking, and substance use during pregnancy: Prevalence, interrelationships, and effects on birth weight. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, 25, 313–320.
- Paarlberg, K. M., Vingerhoets, A. J. J. M., Passchier, J., Heinen, A. G. J. J., Dekker, G. A., & van Geijn, H. P. (1999). Smoking status in pregnancy is associated with daily stressors and low well-being. *Psychology and Health*, 14, 87–96.
- Pickett, K. E., Rathouz, P. J., Kasza, K., Wakschlag, L. S., & Wright, R. (2005). Self-reported smoking, cotinine levels, and patterns of smoking in pregnancy. *Paediatric and Perinatal Epidemiology*, 19, 368–376.
- Pickett, K. E., Wakschlag, L. S., Dai, L., & Leventhal, B. L. (2003). Fluctuations of maternal smoking during pregnancy. *Obstetrics and Gynecology*, 101, 140–147.
- Pickett, K. E., Wakschlag, L. S., Rathouz, P. J., Leventhal, B. L., & Abrams, B. (2002). The working-class context of pregnancy smoking. *Health & Place*, 8, 167–175.
- Pickett, K. E., Wilkinson, R. G., & Wakschlag, L. S. (2009). The psychosocial context of pregnancy smoking and quitting in the Millennium Cohort Study. *Journal of Epidemiology and Community Health*, 63, 474–480.
- Pritchard, C. W. (1994). Depression and smoking in pregnancy in Scotland. *Journal of Epidemiology and Community Health*, 48, 377–382.
- Rice, F., Lewis, A., Harold, G., van den Bree, M., Boivin, J., Hay, D. F., et al. (2007). Agreement between maternal report and antenatal records for a range of pre and peri-natal factors: The influence of maternal and child characteristics. *Early Human Development*, 83, 497–504.

- Sellstrom, E., Arnoldsson, G., Bremberg, S., & Hjern, A. (2008). The neighborhood they live in: Does it matter to women's smoking habits during pregnancy? *Health & Place*, 14, 155–166.
- Selzer, M. L. (1971). The Michigan Alcoholism Screening Test: The quest for a new diagnostic instrument. *American Journal of Psychiatry*, 127, 1653–1658.
- Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry*, 39, 28–38.
- Sherbourne, C. D., & Stewart, A. L. (1991). The MOS social support survey. *Social Science & Medicine*, 32, 705–714.
- Skinner, H. A. (1982). The Drug Abuse Screening Test. *Addictive Behaviors*, 7, 363–371.
- Tager, I. B., Ngo, L., & Hanrahan, J. P. (1995). Maternal smoking during pregnancy. Effects on lung function during the first 18 months of life. *American Journal of Respiratory and Critical Care Medicine*, 152, 977–983.
- Thue, E., Schei, B., & Jacobsen, G. (1995). Psychosocial factors and heavy smoking during pregnancy among parous Scandinavian women. *Scandinavian Journal of Primary Health Care*, 13, 182–187.
- Tomeo, C. A., Rich-Edwards, J. W., Michels, K. B., Berkey, C. S., Hunter, D. J., Frazier, A. L., et al. (1999). Reproducibility and validity of maternal recall of pregnancy-related events. *Epidemiology*, 10, 774–777.
- Wakefield, M., Gillies, P., Graham, H., Madeley, R., & Symonds, M. (1993). Characteristics associated with smoking cessation during pregnancy among working class women. *Addiction*, 88, 1423–1430.
- Wakschlag, L. S., Kistner, E. O., Pine, D. S., Biesecker, G., Pickett, K. E., Skol, A., et al. (2009). Interaction of prenatal exposure to cigarettes and MAOA genotype in pathways to youth antisocial behavior. *Molecular Psychiatry*. Advance online publication March 3, 2009, doi:10.1038/mp.2009.22.
- Wakschlag, L. S., Pickett, K. E., Middlecamp, M. K., Walton, L. L., Tenzer, P., & Leventhal, B. L. (2003). Pregnant smokers who quit, pregnant smokers who don't: Does history of problem behavior make a difference? *Social Science & Medicine*, 56, 2449–2460.
- Weaver, K., Campbell, R., Mermelstein, R., & Wakschlag, L. (2008). Pregnancy smoking in context: The influence of multiple levels of stress. *Nicotine & Tobacco Research*, 10, 1065–1073.
- Yawn, B. P., Suman, V. J., & Jacobsen, S. J. (1998). Maternal recall of distant pregnancy events. *Journal of Clinical Epidemiology*, 51, 399–405.
- Zucker, R. A., Noll, R. B., Ham, H. P., Fitzgerald, H. E., & Sullivan, L. S. (1994). *Assessing antisociality with the Antisocial Behavior Checklist: Reliability and validity studies*. East Lansing: University of Michigan, Ann Arbor and Michigan State University.
- Zuckerman, M. (2002). Zuckerman-Kuhlman Personality Questionnaire (ZKPQ): An alternative five-factorial model. In B. D. Raad & M., Perugini (Eds.), *Big five assessment* (pp. 377–396). Ashland, OH: Hogrefe and Huber.