TECHNICAL PAPERS ON
HEALTH AND BEHAVIOR MEASUREMENT

TECHNICAL PAPER 32

Abortion and Breast Cancer Risk: Fact or Artifact?

Heather G. Miller, James N. Gribble, Leah C. Mazade, and Charles F. Turner

Reference Citation

Miller, H.G., Gribble, J.N., Mazade, L.C., and Turner, C.F. (1999) Abortion and breast cancer risk: Fact or artifact? In Arthur Stone, et al., eds., *Science of Self Report* . Mahwah, N.J.: Lawrence Erlbaum Associates.

CHAPTER

8

The Association Between Self-Reports of Abortion and Breast Cancer Risk: Fact or Artifact

Heather G. Miller
James N. Gribble
Leah C. Mazade
Susan M. Rogers
Charles F. Turner
Research Triangle Institute, Washington, DC

All measurements, whether generated in the laboratory or in the field, are vulnerable to error. Yet all too often, public health policymaking and practice base decisions on such empirical data without taking their methodological limitations fully into account. For example, findings from studies that use small samples of convenience are often given the same weight as the results from larger, population-based samples. And the reporting bias that may afflict self-disclosure of information is not always explored and assessed. Of concern to scientists attending the 1996 National Institutes of Health (NIH) conference and contributing to this volume is the quality of self-reported measurements found in many health studies that address such factors as symptoms, exposures, compliance with treatment, and utilization of health care services.

As researchers have long recognized, several kinds of bias can afflict self-reports. Thus, some understanding of the validity and reliability of measurements is vital to drawing reasonable inferences from them, for methodological artifacts can not only mask important associations between self-reported measures and health outcomes of interest but produce spurious linkages as well. For example, one recent study found no association between self-reported measures of the use of condoms and incident sexually transmitted diseases, or STDs (Turner & Miller, 1997; Zenilman et al., 1995); respondents who reported always using condoms were just as likely to have

an incident STD as people who reported that they never used them. The findings raised troubling questions: Is consistent, correct condom use, a behavior widely considered to prevent transmission of infection, inadequate to protect a person from acquiring an STD? Or is social desirability bias obscuring the protective effect of condoms?

The problem of bias is well documented in the scientific literature but has proved difficult to study. To highlight the importance of exploring bias in self-reports and the challenges involved in such research, this chapter reviews several recent case-control studies that have noted a weak but persistent association between induced abortion and breast cancer. Unfortunately, the findings of two recent studies are being used without qualification in campaigns to discourage women from seeking an abortion. They have appeared in widely disseminated, fear-arousing messages that ignore the uncertainty and inconsistencies in the data. Those problems are carefully referenced in the scientific literature, but advertisements on public transportation in several East Coast cities merely warn that "women who choose abortion suffer more and deadlier breast cancer" (Estrich, 1996). That message delivers misinformation about the severity of the disease in that group of women, and the confusion it may generate among women who are at greatest risk for breast cancer may make them discredit prevention messages about effective strategies, such as mammographic screening.

Sound guidelines for prevention require solid empirical data. Thus, it is crucial to know whether the relationship between induced abortion and breast cancer is real or spurious—that is, an artifact resulting from bias in self-reports of abortion. In this chapter, we examine several potential sources of bias and the quality of existing self-reported data from case-control studies of abortion and breast cancer. Such studies play an important role in the development of public health policies that can affect the lives of millions of women. The scale of that impact argues strongly for improving not only our understanding of bias in key measures but ultimately the quality of those measures.

INTERVIEW CONDITIONS AND BIAS IN SELF-REPORTS OF SENSITIVE BEHAVIORS

Measurements of sensitive behaviors contain many potential sources of bias.[1] Chapter 7, this volume, notes several, including complexities and

[1]Other biases may exist in case-control studies linking abortion and breast cancer besides those discussed in this section. Bias in ascertaining cases, which may result from differential surveillance, diagnosis, or referral, is a well-recognized problem in case-control studies. Case-control research is also vulnerable to biases associated with nonparticipation. Women who have already been identified as cases may be more willing than controls to participate in research studies—with the result that recruitment rates can differ across the two groups.

problems associated with cognitive processing, the social organization of the behavior of interest, task requirements to provide the data, and management of the threat associated with some questions on sensitive or illegal behaviors. Certainly, those sources of bias could be found in measures of abortion as well. Recall bias or difficulties with cognitive processing can be particularly problematic when respondents are asked to report on a lengthy retrospective period. And generally, a diagnosis of breast cancer occurs later in life; it may thus be quite distant from hypothesized causal factors, including abortion. In addition, the task of providing specific information on the timing and outcomes of all conceptions can be enormously complex, especially for women who have suffered multiple miscarriages.

A further consideration is that the social organization of abortion has changed dramatically during the lifetimes of women who are currently diagnosed with breast cancer. What was once an illegal medical procedure no longer carries legal sanctions—yet it remains controversial. A number of psychological and social factors come into play when a woman is asked about a behavior that was once illegal and continues to carry significant stigma.

In recent years, evidence has been building to confirm that the way questionnaires are administered during a survey can affect the quality of self-reported measures of sensitive behaviors. In this chapter, we explore whether the increased privacy afforded by self-administered questionnaires (SAQs) decreases the threat associated with questions about abortion. We also consider whether there are ways to manage the costs associated with SAQs; that is, the task requirements and cognitive complexities inherent in filling out a questionnaire without the assistance of an interviewer.

Paper-and-Pencil Self-Administered Questionnaires

Interviewer-administered questionnaires (IAQs) that contain items on sensitive or illegal behaviors may present problems for both the interviewer and the interviewee. Asking or answering questions, for example, about a respondent's recent history of STDs or date rape could conceivably generate anxiety for both individuals in the interview—for the interviewer, who might worry about losing the case, and for the interviewee, who may fear a "loss of face." Historically, surveys have attempted to deal with the potential for such discomfort by using paper-and-pencil self-administered questionnaires (PAPI SAQs) to increase privacy for both parties.

Data from surveys of several sensitive behaviors, including drug use, sexual behavior, and induced abortion (see the later discussion), support the notion that increased privacy during the interview results in increased reporting of those behaviors. A 1990 field test of the National Household Survey on Drug Abuse (NHSDA) included a methodological experiment that randomized subjects to either PAPI SAQs or IAQs for sensitive portions of

the questionnaire. Overall, PAPI SAQs yielded higher estimates of illicit drug use than did IAQs (Turner, Lessler, & Devore, 1992). However, the relative advantage of PAPI SAQs appeared to be a direct function of the sensitivity of the behavior being reported. Thus, the prevalence of recent cocaine use among NHSDA participants assigned to the PAPI SAQ mode was 2.4 times higher than the prevalence among participants in the IAQ mode. The interview mode effect was less pronounced for reports of marijuana use and was almost nonexistent for reports of alcohol use by adults.[2] The National Longitudinal Survey of Labor Market Experience, Youth Cohort (NLS–Y), also found that PAPI SAQs yielded more frequent reports of illicit drug use than did IAQs (Shober, Fe Caces, Pergamit, & Branden, 1992), as have other such surveys (see, for example, Aquilino, 1994; Gfroerer & Hughes, 1992).

Although few studies have investigated the effect of the mode of interview in surveys of self-reported sexual behavior (Catania, McDermott, & Pollack, 1986), those that have report a mode effect for some but not all sexual behaviors. For example, Millstein and Irwin (1983) found that girls completing IAQs reported significantly lower levels of masturbation and vaginal intercourse (25% and 63%, respectively) compared with girls assigned to either one of the SAQ formats (38% and 74%, respectively).

The increased reports of sensitive behavior associated with more privacy during the interviewing process seem to be a stable finding, but do such reports actually constitute a decrease in reporting bias? In general, survey methodologists believe that biases in the reporting of illicit or stigmatized behavior in general population surveys result in a net negative bias in estimates of the prevalence of those behaviors in the population. The direction of the bias is negative because the number of survey respondents who deny engaging in stigmatized or sensitive behavior in which they have, in fact, engaged is expected to be larger than the number who falsely report behavior in which they have not engaged (Bradburn et al., 1979; Catania, Gibson, Chitwood, & Coates, 1990; Miller, Turner, & Moses, 1990; Turner, Lessler, & Gfroerer, 1992). Thus, researchers believe that higher levels of such reports reflect a reduction in reporting bias and thereby an increase in the accuracy of the measurements.

Concerns about the quality of data gathered in a survey would argue for using PAPI SAQs, which are relatively easy and inexpensive to produce. Nevertheless, their limitations should not be overlooked. Extensive use of contingent questioning—that is, branching or skip patterns—may not be possible in a PAPI SAQ because some respondents have trouble following the complex instructions required to make their way through a self-admin-

[2]A mode effect was found for 12- to 17-year-olds—a group for whom the use of alcohol is illicit. For that group, estimates of the prevalence of recent alcohol use were 1.4 times greater when PAPI SAQs were used.

istered form (Lessler & Holt, 1987). Moreover, according to the National Center for Education Statistics (1993), the reading skills of a sizable segment of the U.S. population are limited, which means that some proportion of survey respondents will not be able to complete these forms by themselves. Because literacy is correlated with education and other indicators of socioeconomic status, the resulting measurements from surveys that cannot hold privacy constant across all subjects may be afflicted with other biases as well. For example, surveys that use PAPI SAQs for sensitive portions of the questionnaire run the risk of producing differences in response distributions if participants of limited literacy cannot be randomized to that mode. Moreover, some investigators have noted higher rates of item nonresponse with PAPI SAQs in comparison to IAQs (Cox, Witt, Traccarella, & Perez-Michael, 1992; Fay, Turner, Klassen, & Gagnon, 1989; Rogers & Turner, 1991; Turner, Miller, & Moses, 1989). And among participants who do respond to all items, the number who provide logically inconsistent answers on PAPI SAQs can be substantial (see, for example, Cox et al., 1992, and Smith, 1992).

Audio Computer-Assisted Self-Interviewing

Until recently, the problems just noted were for the most part unavoidable but advances in interviewing technology now offer some hope of overcoming them. Specifically, the use of audio computer-assisted self-interviewing (audio-CASI) appears to reduce some of the measurement biases that have been problematic in prior surveys of sensitive behavior (Cooley, Turner, O'Reilly, Allen, & Paddock, 1996; Johnston, 1992; O'Reilly & Turner, 1992; O'Reilly, Hubbard, Lessler, Biemer, & Turner, 1994).

With audio-CASI, respondents use portable laptop computers and listen to questions through headphones; they enter their answers by pressing labeled keys. The system does not rely on synthesized voices, and the recorded audio component is equivalent to a high-quality tape recording. In addition, there are no significant delays in playing back the audio-delivered questions. Unlike some earlier efforts (e.g., Camburn, Cynamon, & Harel, 1991), audio-CASI is capable of executing skip patterns, checking for out-of-range responses and inconsistencies across similar questions, and generating data files. Every respondent hears the questions asked in the same neutral manner, regardless of the question's substance or the response it elicits from a participant. The technology can be used with any respondent who can hear and speak and does not require literacy in any language (see, for example, Hendershot, Rogers, Thornberry, Miller, & Turner, 1996; Turner, Rogers, Hendershot, Miller, & Thornberry, 1996).

Early studies hypothesized that the use of audio-CASI would reduce reporting bias by increasing privacy during the interview for all respondents, even those of limited literacy. The results of several large studies (Duffer,

Lessler, Weeks, & Mosher, 1996; Turner, Ku, Sonenstein, & Pleck, 1996; Turner, Lessler, & Devore, 1992; Turner, Miller, & Rogers, 1997) that contained an audio-CASI component, including studies on drug use and AIDS-related sexual behaviors, indicate just such a reduction.

Mail and Telephone Interviews

Because studies that collect interview data from large population-based samples can be very expensive, many researchers have used mailed questionnaires and telephone interviews when logistic and economic constraints precluded face-to-face interviews. (See, for example, two recent telephone surveys: the National AIDS Behavioral Survey, Catania et al., 1992, and a French survey of AIDS-related sexual behavior, ACSF Investigators, 1992.) Yet each mode of data collection has its own set of limitations. Surveys that are conducted by mail afford participants privacy but often have poor response rates. Telephone surveys generally have better response rates but are not similarly private because they require participants to report information to another human being, albeit over the phone rather than face to face. The problem of privacy in telephone interviewing and the apparent success of audio-CASI in reducing bias in self-reports led scientists at Research Triangle Institute to explore the potential of audio-CASI for use in telephone interviewing.

A New Telephone Interviewing Technology

Telephone audio-CASI, or T-ACASI, is built on the audio-CASI platform but uses a touchtone telephone instead of a laptop computer. It thus offers the same advantages as audio-CASI in terms of survey administration, although operationally it replicates many of the procedures of standard telephone interviewing.[3]

Results from two T-ACASI pilot studies (Miller et al., 1997; Turner, Miller, Smith, Cooley, & Rogers, 1996) indicate that the system is stable and easy to use and respondents prefer it to a live interviewer when sensitive questions are being asked. Moreover, a study that compared data from standard telephone interviewing with data collected by T-ACASI found that T-ACASI interviews produced higher rates of prevalence of several sensitive behaviors, including anal intercourse, and lower rates of normative behaviors, such as condom use (Turner, Miller et al., 1996). Furthermore, T-ACASI

[3]In a T-ACASI interview, a human telephone interviewer calls and recruits an eligible respondent. After eliciting consent and collecting data on nonsensitive items, the interviewer transfers the call to the T-ACASI system, which administers the sensitive part of the questionnaire.

participants were two times more likely than respondents assigned to the live interviewer to report never using a condom.

The potential value of T-ACASI in reducing bias in self-reports has yet to be fully explored, although several studies are under way. Its application to measures of abortion, a major topic of interest here, may offer a way to investigate whether a link indeed exists between induced abortion and breast cancer—or whether the association observed in other case-control studies reflects differential reporting of cases versus controls. The next section reviews the evidence from case-control studies on the role that reporting bias may play in that research.

EVIDENCE OF A LINK BETWEEN ABORTION AND BREAST CANCER

Two recent case-control studies (Daling, Malone, Voigt, White, & Weiss, 1994; Newcomb et al., 1996) reported a modest but statistically significant association between breast cancer and induced abortion. Because of the intensity of the abortion debate in this country, the finding was quickly seized on and widely publicized. Left behind was a carefully balanced review of existing data provided by the Daling team, the caveats included by Newcomb and her colleagues, and statements issued by researchers from the National Cancer Institute and other scientific authorities advising caution. All of those qualifying comments went virtually unheeded in the lay press, written off, perhaps, as academic methodological hair-splitting. The National Cancer Institute (1994), in its presentation and review of the findings of Daling and coworkers, summarized some of the methodological concerns that make it difficult to draw inferences from the studies, noting that small relative risks, such as those reported by Daling and by Newcomb, are "usually difficult to interpret. Such increases may be due to chance, statistical bias, or effects of confounding factors that are sometimes not evident" (p. 2).

We reviewed the epidemiological literature for other studies of the relationship between breast cancer and abortion (see Table 8.1) and found no definitive answers. Abundant and solid empirical evidence supports the premise that full-term pregnancy reduces the risk of breast cancer (e.g., Kelsey, Gammon, & John, 1993; MacMahon et al., 1970). But the picture offered by studies of the relationship between pregnancies that do not go to full term and the risk of breast cancer is much less clear. Some investigators report no significant increase in the risk of breast cancer with either spontaneous or induced abortion (Adami, Bergstrom, Lund, & Meirik, 1990; Brinton, Hoover, & Fraumeni, 1983; La Vecchia et al., 1987; Lipworth, Katsouyanni, Ekbom, Michels, & Trichopoulos, 1995; Parazzini, La Vecchia, & Negri, 1991; Rosenberg et al., 1988). Others have found increased risk asso-

TABLE 8.1
Selected Major Case-Control Studies of Breast Cancer and Abortion

Citation	*N*	*Response Rate*	*Mode*	*Population*	*RR Induced Abortion (95% CI)*	*RR Spontaneous Abortion (95% CI)*
Newcomb et al., 1996	Cases: 6,888 Controls: 9,529	Case: 81% Control: 84%	IAQ (phone)	U.S. women < 75 years old	1.23 (1.00–1.51)	1.11 (1.02–1.20)
Lipworth et al., 1995	Cases: 820 Controls: 1,548	Case: 94% Control: 94%	IAQ	Greek women (nulliparous)	0.98 (0.56–1.73)	1.17 (0.64–2.13)
				Greek women (parous)	0.99 (0.56–1.74)	0.61 (0.33–1.14)
Daling et al., 1994	Cases: 845 Controls: 961	Case: 84% Control: 74%	IAQ	U.S. White females < 46 years old	1.5 (1.2–1.9)	0.9 (0.7–1.2)
Parazzini et al., 1991	Cases: 2,394 Controls: 2,218	Case: 97% Control: 98%	IAQ	Italian females < 75 years old	0.9 (0.7–1.1)	0.8 (0.7–1.1)
Adami et al., 1990	Cases: 422 Controls: 527	Case: 89% Control: 81%	IAQ	Scandinavian women < 45 years old	1.3* (0.6–3.0)	1.3* (0.7–2.6)
Ewertz and Duffy, 1988	Cases: 1,486 Controls: 1,336	Case: 88% Control: 79%	SAQ (mail)	Danish women < 70 years old (no full-term pregnancy)	3.85 (1.08–13.6)	2.63 (0.83–8.32)
				Danish women < 70 years old (full-term pregnancy)	N.S.**	N.S.**

Rosenberg et al., 1988	Cases: 3,200 Controls: 4,844	***	IAQ	U.S. women < 70 years old (nulliparous)	1.3 (0.9–2.2)	0.9 (0.5–1.5)
				U.S. women < 70 years old (parous)	1.2 (0.9–1.6)	0.9 (0.8–1.0)
La Vecchia et al., 1987	Cases: 1,108 Controls: 1,281	Case: 98% Control: 98%	IAQ	Italian women < 75 years old	0.7 (0.5–1.1)	0.9 (0.7–1.3)
Hirohata et al., 1985	Cases: 212 Controls: 424	Case: 99% Controls: 87%	IAQ	Japanese women	1.19 (0.7–1.9)	1.53 (1.01–2.3)
Brinton et al., 1983	Cases: 1,362 Controls: 1,250	Case: 86% Control: 74%	IAQ	U.S. White females (before live birth)	1.34 (0.3–5.6)	1.09 (0.8–1.5)
				U.S. White females (after live birth)	0.89 (0.4–2.0)	1.2 (0.9–1.6)

Note. IAQ = interviewer-administered questionnaire; SAQ = self-administered questionnaire.

*Adjusted risk ratio (RR) for women reporting 2 or more abortions (spontaneous or induced).

**N.S. (no significant association).

***The authors do not provide response rates for cases and controls. They do state that 5% of patients (which would include both cases and controls) declined participation.

ciated with both types of abortion (Howe, Senie, Bzduch, & Herzfeld, 1989; Newcomb et al., 1996). Some studies report increased risk associated with induced but not spontaneous abortion (Daling et al., 1994; Ewertz & Duffy, 1988), and at least one has found increased risk only with spontaneous abortion (Hirohata et al., 1985). Several large cohort studies have also looked at the association between breast cancer and abortion, but again, their data do not allow definitive statements about any such link (see Calle et al., 1995; Hadjimichael, Boyle, & Meigs, 1986; Kvale, Heuch, & Eide, 1987; Lindefors-Harris, Eklund, Meirik, Rutqvist, & Wiklund, 1989).

The variation in findings across case-control studies may reflect small samples, differences in the populations being sampled, or differences in design and analysis, including controlling for different covariates. Case-control studies—like all experimental designs—have strengths and weaknesses that are known to affect the quality of the data they produce (see, for example, Breslow & Day, 1980; Sackett, 1979; Schlesselman, 1982). When such studies produce small risk ratios or findings at tenuous levels of significance, bias as an alternative explanation of those findings becomes more plausible. Because the size of the effect in much of the research is small, several investigators have wondered whether the abortion–breast cancer association might be an artifact of reporting bias (Daling et al., 1994; Gammon, Bertin, & Terry, 1996; Henshaw, 1996; Lindefors-Harris et al., 1989; Lindefors-Harris, Eklund, Adami, & Meirik, 1991; Lipworth et al., 1995; Michels, Hsieh, Trichopoulos, & Willett, 1995; Newcomb et al., 1996). Evidence from other large-scale studies of women that included items on abortion supports the notion that measurements of this sensitive topic are vulnerable to reporting bias.

Reporting of Abortions

There are two sources of data on induced abortions in the United States: clinicians who perform abortions and the women who undergo them. Surveillance data from clinicians are collected by the Centers for Disease Control and Prevention (CDC); self-reported data come from health surveys conducted among women of childbearing age.[4] Comparing surveillance data with self-reported data provides estimates of some of the bias in the self-reported measures.

According to CDC data, 1,429,577 abortions were performed in 1990 (Koonin, Smith, & Ramick, 1993). Overall, the majority of women who underwent abortion in 1990 were young, and almost half did not report a previous live birth. In an independent survey of abortion providers, Henshaw, Koonin,

[4]That is, the National Survey of Family Growth (NSFG), the National Surveys of Young Women (NSYW), and the National Longitudinal Survey of Labor Market Experience, Youth Cohort.

and Smith (1991) found few induced abortions among women older than 40 but substantial differences by race: The abortion rate among non-White females was 2.7 times the rate among White females.

Self-reported data confirm some of the patterns seen in surveillance data but show major dissimilarities in other areas. Like the surveillance data, self-reports show that younger women were more likely than older women to report an abortion and that abortions were most frequently reported by women who reported no previous live births (Jones & Forrest, 1992; Miller et al., 1997). The completeness of the self-reporting, however, differs substantially from the CDC data. Overall, the abortions reported by women taking part in health surveys accounted for only 35% to 40% of the abortions enumerated in surveillance data (Jones & Forrest, 1992). Moreover, completeness of reporting appears to vary by race: A slightly greater proportion of Black females versus White females reported a lifetime history of abortion. But the actual racial difference appears to be even greater. Comparing data from the 1988 wave of the NSFG with abortion surveillance data, Jones and Forrest found that White females reported 38% of the abortions noted in surveillance records, whereas non-White females reported 27% of such abortions.[5]

Reporting bias may also vary by the conditions under which data are collected. Under standard interview conditions in the NSFG, interviewers asked respondents about their abortion history during a face-to-face interview using IAQs. But in an experiment embedded in the 1988 wave of the NSFG, respondents were also given a PAPI SAQ to offer them a second, private opportunity to report past abortions. Use of the PAPI SAQs increased women's reporting of abortions from 39% to 71% of the level reported by abortion providers (Jones & Forrest, 1992). The impact of the interview mode may also vary by race. Data from the experiment showed that the level of completeness of abortion reporting among White females increased from 46% in the IAQ mode to 74% in the PAPI SAQ mode (Jones & Forrest, 1992). However, the increase among Black women was greater, rising from 26% in the IAQ mode to 67% in the PAPI SAQ mode.

In 1994, RTI conducted a pilot test for the 1995 wave of the NSFG that included a mode experiment. In the pilot study, interviewers asked women about their sexual history, including any abortions they might have had. At the end of the IAQ portion of the interview, women were reinterviewed using audio-CASI.[6] The proportion of women who reported abortions in the private

[5]Comparing data from the NLS-Y with surveillance data indicated approximately the same rate of underreporting, with White females reporting 45% of abortions, Black females 27%, and Hispanic females 19%.

[6]The audio-CASI interview included some items that the interviewer had asked before as well as a second section of new questions about such sensitive behaviors as injecting drugs, needle sharing, and same-gender sex.

audio-CASI condition (23.6%) was 1.4 times greater than the proportion who reported them in the less-private IAQ mode (17.3%; Duffer et al., 1996). Black women were more than twice as likely as White women (31% versus 14%) to report more abortions or sexual partners in the past 12 months in the audio-CASI mode (Kinsey, Thornberry, Carson, & Duffer, 1995).

The now completed 1995 wave of the NSFG replicated the pilot study's mode experiment for sensitive items, including abortion. All of the women who completed the NSFG's face-to-face IAQ (10,847 respondents) were asked to complete a smaller, self-administered reinterview using RTI's audio-CASI system. Preliminary analyses of the data from the 1995 wave (Miller et al., 1997) appear to confirm the findings of the mode experiment from the 1988 wave (see Table 8.2). That is, among sexually active women, the number who reported a history of abortion in the more private audio-CASI mode was greater than in the interviewer-administered mode. Of the women who did not report an abortion in the original IAQ mode, 4.5% reported one or more abortions in the audio-CASI mode.[7] Among women reporting one abortion in the IAQ mode, 5.8% reported two or more abortions in the audio-CASI mode.[8] Most of the discrepancies in reports of abortion across modes occurred in only one direction. Thus, among women reporting two or more abortions in the IAQ mode, only 0.4% reported no abortions in the audio-CASI mode. As the earlier wave found, the mode effect was greater among Black women than among Whites: 7.3% of Black females who reported no abortions in the IAQ mode reported one or more abortions when using audio-CASI; the comparable percentage for White participants was 4.2%. In addition, 10.3% of Black females reporting one abortion in the IAQ mode reported two or more abortions in the audio-CASI mode, compared with 5.0% of White women. Overall, the odds of reporting an abortion were approximately 1.3 times greater when information was collected using the audio-CASI technology compared with the IAQ.

At least two arguments support the hypothesis that data from the NSFG experiment provide a lower-bound estimate of the effect of the interview mode on measures of abortion. First, the design of the NSFG experiment required that all women complete the IAQ before the audio-CASI reinterview. A woman who wanted to answer the question about abortion more honestly

---

[7]Of 9,674 sexually active respondents who were interviewed in the IAQ mode, 2,121 reported no history of pregnancy, and consequently interviewers did not administer the questions concerning abortion. In this analysis, those women are considered as reporting no abortions. All women, regardless of pregnancy history, received questions on abortion in the audio-CASI mode.

[8]All percentages are weighted to account for variation in the probability of selection and nonresponse. (Weighting allows for comparison of these estimates with 1995 census projections.)

TABLE 8.2
Comparison of Abortion Reporting by Interview Mode
Among Sexually Active Women: 1995 NSFG

	SUBSEQUENT REPORT IN A-CASI			
ORIGINAL IAQ REPORT	*No Abortions* %	*1 Abortion* %	*2+ Abortion* %	*Unweighted N*
All Women				
No Abortions*	95.4	3.5	1.0	*7,827*
1 Abortion	1.8	92.5	5.8	*1,265*
2+ Abortions	0.4	2.2	97.4	*582*
White Women				
No Abortions	95.8	3.3	0.9	*5,675*
1 Abortion	1.1	93.9	5.0	*850*
2+ Abortions	0.6	2.5	96.9	*343*
Black Women				
No Abortions	92.8	5.3	2.0	*1,742*
1 Abortion	4.4	85.3	10.3	*355*
2+ Abortions	0.0	1.8	98.2	*217*

Note. The analyses exclude 230 cases with missing data.

*In the 1995 IAQ, women were asked to report pregnancy histories for up to 15 pregnancies, and abortion was listed as one of the possible pregnancy outcomes. Therefore, women who reported never being pregnant were not questioned about abortion. Included in these estimates are 54 women (16 Black, 37 White, 1 other race) who reported in the IAQ mode that they had never been pregnant but who reported at least one abortion in the audio-CASI mode.

(i.e., differently) in the audio-CASI mode may have found herself in a quandary. Answering more honestly in the audio-CASI mode would reveal her dishonesty on the IAQ. To prevent that, some women may have chosen to be consistent rather than honest. Second, the NSFG sample included only women between 15 and 44 years of age. The older cohort of women who could report an illegal abortion (i.e., one that occurred prior to 1971) was not represented in the sample.

Despite clear evidence of bias in self-reported measures of abortion, researchers have few alternatives to self-disclosure for gathering retrospective data on a controversial procedure. Improving our understanding of reporting bias in measures of abortion becomes critical when abortion is being considered as a potential risk factor for a serious medical condition or disease such as breast cancer. The next section reviews the evidence available from case-control studies of abortion and breast cancer risk and how reporting bias may affect inferences drawn from those results.

Potential Sources of Bias in Case-Control Studies of Abortion and Breast Cancer

Recall bias, or the differential ability to remember a specific piece of information, is not unique to case-control studies, but in case-control studies, it may have an additional dimension. Compared with controls, cases may have a differential capacity to remember their medical history. During diagnosis and treatment of their disease, women who have recently been diagnosed with breast cancer have been asked questions about many possible risk factors. They may thus be cognitively primed to remember historical events of interest to researchers. That argument, however, does not seem very compelling for abortion. Given the psychological and physical trauma surrounding induced abortion, it is difficult to believe that women who have ever had an abortion will forget it. What is more plausible is that cases may be more motivated to provide a complete medical history because they may feel that they have much to gain from cooperating with people who are trying to understand their disease. Prospective studies, which are less vulnerable to recall bias, have generally not found an increase in the risk of breast cancer among women who report abortions (Calle et al., 1995; Kvale et al., 1987; Lindefors-Harris et al., 1989; Sellers et al., 1993).

Several investigators have hypothesized that differential reporting of abortion by cases and controls may play a role in findings of a link between abortion and breast cancer. They speculate that women without the disease may have less compelling reasons for participating in research and thus may be less inclined to report prior abortions. A study conducted in Sweden lends some support to that hypothesis (Lindefors-Harris et al., 1991); in comparing registry reports of induced abortions with interview data, the investigators found greater underreporting of induced abortion among healthy controls than among women with newly diagnosed breast cancer. The study compared interview data with registry data that covered legally induced abortions in Sweden for the 1966 to 1974 period. Interviews were conducted with 317 women with a history of breast cancer and 512 female controls; the study used IAQs to collect retrospective data on induced and spontaneous abortions. Cases and controls were identified by their national registration number, which allowed researchers to link data from study participants with data from national health registries, including the abortion registry, and assess their concordance. Among cases, 26 women admitted to abortion, but the registry only identified 24 cases. Among controls, 44 women admitted to abortion, but the registry identified 59 controls.

Finally, there are questions about social desirability bias, which can result when subjects feel constrained to report the socially acceptable response. In trying to probe differential bias in cases versus controls, Newcomb et al. (1996) looked at reports of induced abortion by legal status (that is, before

1973, when abortions were illegal, compared with after 1973, when they were legal). They found that the risk ratio was greater for reports of illegal induced abortion than for reports of legal abortion, which suggests that cases were more likely than controls to report those events when they occurred under conditions of greater social and legal constraint.

CONCLUSION

The scientific literature leaves little doubt that self-reported measures of sensitive behaviors, including abortion, are vulnerable to bias. Moreover, findings from case-control studies indicate that the magnitude of that bias may not be constant across all subjects. The question that remains is what to do about the problem of bias and flawed estimates of critical variables. The availability of new automated interview technology holds promise for further methodologic research to improve our understanding of bias in sensitive data gathered in future case-control studies.

Although methodologic research is often viewed as esoteric, myopic, or worse, the pursuit of methodological research on the quality of self-reported data on abortion holds promise of some very concrete benefits. For example, improved measures of abortion could improve our capability to evaluate new contraceptive technologies and to interpret abortion surveillance data. Moreover, efforts to bring improved data to address the reported link between breast cancer and abortion potentially could affect a broader range of individuals, all of whom are currently operating in a politically charged atmosphere. Individual women, policy makers, service providers, and professionals crafting prevention messages all need to know whether the relationship between abortion and breast cancer is real or a methodologic artifact. To be able to speak with more confidence about the association between breast cancer and abortion would certainly be a step forward.

ACKNOWLEDGMENTS

This chapter draws on a presentation (Turner and Miller) at the 1996 Conference on the Science of Self-Reports, sponsored by the National Institutes of Health, and a related research proposal (Miller, Turner, Helzlsouer, Zenilman, & Newcomb, 1997). Coinvestigators Polly Newcomb, Kathy Helzlsouer, and Jonathan Zenilman contributed to the latter effort, but they are not responsible for any errors that may appear in this chapter. Its preparation was supported by grant RO1-HD/AG31067-03 from the National Institute of Child Health and Human Development and the National Institute on Aging, and grant R01-MH56318-01 from the National Institute of Mental Health.

REFERENCES

ACSF Investigators. (1992). Analysis of sexual behavior in France: A comparison between two modes of investigation: Telephone survey and face-to-face survey. *AIDS, 6*, 315–323.

Adami, H. O., Bergstrom, R., Lund, E., & Meirik, O. (1990). Absence of association between reproductive variables and the risk of breast cancer in young women in Sweden and Norway. *British Journal of Cancer, 62*, 122–126.

Aquilino, W. S. (1994). Interview mode effects in surveys of drug and alcohol use. *Public Opinion Quarterly, 58*, 210–240.

Bradburn, N. M., Sudman, S., Blair, S., Locander, W., Miles, C., Singer, E., & Stocking, C. (1979). *Improving interview method and questionnaire design*. San Francisco: Jossey Bass.

Breslow, N. E., & Day, N. E. (1980). *Statistical methods in cancer research. Vol. 1: The analysis of case-control studies*. Lyon, France: International Agency for Research on Cancer.

Brinton, L. A., Hoover, R., & Fraumeni, J. F. (1983). Reproductive factors in the aetiology of breast cancer. *British Journal of Cancer, 47*, 757–782.

Calle, E. E., Merris, C. A., Wingo, P. A., Thun, M. J., Rodriguez, C., & Heath, C. W., Jr. (1995). Spontaneous abortion and risk of fatal breast cancer in a prospective cohort of United States women. *Cancer Causes and Control, 6*, 460–468.

Camburn, D., Cynamon, D., & Harel, Y. (1991, May). *The use of audio tapes and written questionnaires to ask sensitive questions during household interviews*. Presentation to the National Field Technologies Conference, San Diego, CA.

Catania, J. A., Coates, T. J., Stall, R., Turner, H., Peterson, J., Hearst, N., Dolcini, M. M., Hudes, E., Gagnon, J., Wiley, & Groves, R. (1992). Prevalence of AIDS-related risk factors and condom use in the United States. *Science, 258*, 1101–1106.

Catania, J. A., Gibson, D. R., Chitwood, D. D., & Coates, T. J. (1990). Methodological problems in AIDS behavioral research: Influences on measurement error and participation bias in studies of sexual behavior. *Psychology Bulletin, 108*, 339–362.

Catania, J., McDermott, L., & Pollack, L. (1986). Questionnaire response bias and face-to-face interview sample bias in sexuality research. *Journal of Sex Research, 22*, 52–72.

Cooley, P. C., Turner, C. F., O'Reilly, J. M., Allen, D. R., & Paddock, R. E. (1996). Audio-CASI: Hardware and software considerations in adding sound to a computer-assisted interviewing system. *Social Science Computer Review, 14*, 197–204.

Cox, B., Witt, M., Traccarella, M., & Perez-Michael, A. (1992). Inconsistent reporting of drug use in 1988. In C. F. Turner, J. T. Lessler, & J. D. Gfroerer (Eds.), *Survey measurement of drug use* (pp. 109–153). DHHS Publication No. 92-1929. Washington, DC: Government Printing Office.

Daling, J. R., Malone, K. E., Voigt, L. F., White, E., & Weiss, N. S. (1994). Risk of breast cancer among young women: Relationship to induced abortion. *Journal of the National Cancer Institute, 86*, 1584–1592.

Duffer, A. P., Lessler, J. T., Weeks, M. F., & Mosher, W. D. (1996). Impact of incentives and interviewing modes: Results from the National Survey of Family Growth Cycle V pretest. In R. Warnecke (Ed.), *Health survey research methods: Conference proceedings* (pp. 147–151). DHHS Pub No. (PHS) 96-1013. Hyattsville, MD: National Center for Health Statistics.

Estrich, S. (1996, February 1). Right plays politics with breast cancer. *USA Today*, p. A9.

Ewertz, M., & Duffy, S. W. (1988). Risk of breast cancer in relation to reproductive factors in Denmark. *British Journal of Cancer, 58*, 99–104.

Fay, R., Turner, C. F., Klassen, A., & Gagnon, J. (1989). Prevalence and patterns of same-gender sexual contact among men. *Science, 243*, 338–348.

Gammon, M. D., Bertin, J. E., & Terry, M. B. (1996). Abortion and the risk of breast cancer: Is there a believable association? *Journal of the American Medical Association, 275*, 321–322.

Gfroerer, J. C., & Hughes, A. L. (1992). Collecting data on illicit drug use by phone. In C. F. Turner, J. T. Lessler, & J. D. Gfroerer (Eds.), *Survey measurement of drug use* (pp. 277–295). DHHS Publication No. 92-1929. Washington, DC: Government Printing Office.

Hadjimichael, O. C., Boyle, C. A., & Meigs, J. W. (1986). Abortion before first live birth and risk of breast cancer. *British Journal of Cancer, 53*, 281–284.

Hendershot, T. P., Rogers, S. M., Thornberry, J. T., Miller, H. G., & Turner, C. F. (1996). Multilingual audio-CASI: Using English-speaking field interviewers to survey elderly Korean households. In R. B. Warnecke (Ed.), *Health survey research methods: Conference proceedings* (pp. 165–169). DHHS Pub. No. (PHS) 96-1013. Hyattsville, MD: National Center for Health Statistics.

Henshaw, S. K. (1996). Pregnancy termination and risk of breast cancer. *Journal of the American Medical Association, 276*, 31.

Henshaw, S. K., Koonin, L. M., & Smith, J. C. (1991). Characteristics of U.S. women having abortions, 1987. *Family Planning Perspectives, 23*, 75–81.

Hirohata, T., Shigematsu, T., Nomura, A. M. Y., Nomura, Y., Horie, A., & Hirohata, I. (1985). Occurrence of breast cancer in relation to diet and reproductive history: A case-control study in Fukuoka, Japan. *National Cancer Institute Monograph, 69*, 187–191.

Howe, H. L., Senie, R. T., Bzduch, H., & Herzfeld, P. (1989). Early abortion and breast cancer risk among women under 40. *International Journal of Epidemiology, 18*, 300–304.

Johnston, G. (1992, January). *Demonstration of computer-administered audio survey technology*. Seminar presented at the National Center for Health Statistics, Hyattsville, MD.

Jones, E. F., & Forrest, J. D. (1992). Underreporting of abortion in surveys of U.S. women: 1976 to 1988. *Demography, 29*, 113–126.

Kelsey, J. L., Gammon, M. D., & John, E. M. (1993). Reproductive factors and breast cancer. *Epidemiology Review, 15*, 36–47.

Kinsey, S., Thornberry, J., Carson, C., & Duffer, A. (1995). Respondent preferences toward audio-CASI and how that affects data quality. In *Proceedings of the joint ASA/AAPOR meeting, 1995*. Washington, DC: American Statistical Society.

Koonin, L. M., Smith, J. C., & Ramick, M. (1993). Abortion surveillance: United States, 1990. *Morbidity and Mortality Weekly Review, 42(SS-6)*, 29–57.

Kvale, G., Heuch, I., & Eide, G. E. (1987). A prospective study of reproductive factors and breast cancer. *American Journal of Epidemiology, 126*, 831–841.

La Vecchia, C., Decarli, A., Parazzini, F., Gentile, A., Negri, E., Cecchetti, G., & Franceschi, S. (1987). General epidemiology of breast cancer in Northern Italy. *International Journal of Epidemiology, 16*, 347–355.

Lessler, J. T., & Holt, M. (1987). Using response protocols to identify problems in the U.S. Census long form. In *Proceedings of the Section on Survey Methods Research, American Statistical Association*, pp. 262–265.

Lindefors-Harris, B. M., Eklund, G., Adami, H. O., & Meirik, O. (1991). Response bias in a case-control study: Analysis utilizing comparative data concerning legal abortions from two independent Swedish studies. *American Journal of Epidemiology, 134*, 1003–1008.

Lindefors-Harris, B. M., Eklund, G., Meirik, O., Rutqvist, L. E., & Wiklund, K. (1989). Risk of cancer of the breast after legal abortion during first trimester: A Swedish registry study. *British Medical Journal, 299*, 1430–1432.

Lipworth, L., Katsouyanni, K., Ekbom, A., Michels, K. B., & Trichopoulos, D. (1995). Abortion and the risk of breast cancer: A case-control study in Greece. *International Journal of Cancer, 61*, 181–184.

MacMahon, B., Cole, P., Lin, T. M., Lowe, C. R., Mirra, A. P., Ravnihar, B., Salber, E. J., Valaoras, V. G., & Yuasa, S. (1970). Age at first birth and breast cancer risk. *Bulletin of the World Health Organization, 43*, 209–221.

Michels, K. B., Hsieh, C. C., Trichopoulos, D., & Willett, W. C. (1995). Abortion and breast cancer risk in seven countries. *Cancer Causes and Control, 6*, 75–82.

Miller, H. G., Turner, C. F., & Moses, L. E. (Eds.). (1990). *AIDS: The second decade*. Washington DC: National Academy Press.

Miller, H. M., Turner, C. F., Helzlsouer, K. J., Zenilman, J. M., & Newcomb, P. A. (1997). Abortion, breast cancer, and reporting bias. Unpublished proposal submitted to the National Institute

of Child Health and Human Development and the National Cancer Institute. Research Triangle Park, NC: Research Triangle Institute.

Millstein, S., & Irwin, C. (1983). Acceptability of computer-acquired sexual histories in adolescent girls. *Journal of Pediatrics, 103*, 815–819.

National Cancer Institute. (1994, December). Abortion and possible risk for breast cancer: Analysis and inconsistencies. In *Cancer Facts*. Bethesda, MD: Author.

National Center for Education Statistics. (1993). *Adult literacy in America: A first look at the results of the National Adult Literacy Survey*. Washington, DC: U.S. Department of Education.

Newcomb, P. A., Storer, B. E., Longnecker, M. P., Mittendorf, R., Greenbert, E. R., & Willett, W. C. (1996). Pregnancy termination in relation to risk of breast cancer. *Journal of the American Medical Association, 275*, 283–287.

O'Reilly, J., Hubbard, M., Lessler, J., Biemer, P., & Turner, C. F. (1994). Audio computer-assisted self-interviewing: New technology for data collection on sensitive issues and special populations. *Journal of Official Statistics, 10*, 197–214.

O'Reilly, J., & Turner, C. F. (1992, March). *Survey interviewing using audio format, computer-assisted technologies*. Presentation to the Washington Statistical Society.

Parazzini, F., La Vecchia, C., & Negri, E. (1991). Spontaneous and induced abortions and risk of breast cancer. *International Journal of Cancer, 48*, 816–820.

Rogers, S. M., & Turner, C. F. (1991). Patterns of same-gender sexual contact among men in the U.S.A., 1970–1990. *Journal of Sex Research, 28*, 491–519.

Rosenberg, L., Palmer, J. R., Kaufman, D. W., Strom, B. L., Schottenfeld, D., & Shapiro, S. (1988). Breast cancer in relation to the occurrence and time of induced and spontaneous abortion. *American Journal of Epidemiology, 127*, 981–989.

Sackett, D. L. (1979). Bias in analytic research. *Journal of Chronic Diseases, 32*, 51–68.

Schlesselman, J. J. (1982). *Case control studies: Design, conduct, analysis*. New York: Oxford University Press.

Sellers, T. A., Potter, J. D., Severson, R. K., Bostick, R. M., Nelson, C. L., Kushi, L. H., & Folsom, A. R. (1993). Difficulty becoming pregnant and family history as interactive risk factors for post-menopausal breast cancer: The Iowa Women's Health Study. *Cancer Causes and Control, 4*, 21–28.

Shober, S. E., Fe Caces, M., Pergamit, M. R., & Branden, L. (1992). Effect of mode of administration on reporting in the National Longitudinal Survey. In C. F. Turner, J. T. Lessler, & J. D. Gfroerer (Eds.), *Survey measurement of drug use* (pp. 267–276). DHHS Publication No. 92-1929. Washington, DC: Government Printing Office.

Smith, T. (1992). A methodological analysis of the sexual behavior questions on the General Social surveys. *Journal of Official Statistics, 8*, 309–325.

Turner, C. F., Ku, L., Sonenstein, F. L., & Pleck, J. H. (1996). Impact of Audio-CASI on bias in reporting of male-male sexual contacts. In R. B. Warnecke (Ed.), *Health survey research methods: Conference proceedings* (pp. 171–176). DHHS Pub. No. (PHS) 96-1013. Hyattsville, MD: National Center for Health Statistics.

Turner, C. F., Lessler, J., & Devore, J. (1992). Effects of mode of administration and wording on reporting of drug use. In C. F. Turner, J. T. Lessler, & J. D. Gfroerer (Eds.), *Survey measurement of drug use: Methodological issues* (pp. 177–219). DHHS Pub. No. 92-1929, Washington, DC: Government Printing Office.

Turner, C. F., Lessler, J., & Gfroerer, J. (1992). Future directions for research and practice. In C. F. Turner, J. T. Lessler, & J. D. Gfroerer (Eds.), *Survey measurement of drug use: Methodological issues* (pp. 299–306). DHHS Pub. No. 92-1929, Washington, DC: Government Printing Office.

Turner, C. F., & Miller, H. G. (1997). Zenilman's anomaly reconsidered: Fallible reports, *ceteris paribus*, fragile condoms, and other hypotheses. *Sexually Transmitted Diseases, 24*, 522–527.

Turner, C. F., Miller, H. G., & Moses, L. E. (Eds.). (1989). *AIDS, sexual behavior, and intravenous drug use*. Washington, DC: National Academy Press.

Turner, C. F., Miller, H. G., & Rogers, S. M. (1997). Survey measurement of sexual behaviors: Problems and progress. In J. Bancroft (Ed.), *Researching sexual behavior* (pp. 37–60). Bloomington: Indiana University Press.

Turner, C. F., Miller, H. G., Smith, T. K., Cooley, P. C., & Rogers, S. M. (1996). Telephone audio computer-assisted self-interviewing (T-ACASI) and survey measurements of sensitive behaviors: Preliminary results. In R. Banks, J. Fairgrieve, L. Gerrard, T. Orchard, C. Payne, & A. Westlake (Eds.), *Survey and statistical computing 1996* (pp. 121–130). Chesham, Bucks, UK: Association for Survey Computing.

Turner, C. F., Rogers, S. M., Hendershot, T., Miller, H., & Thornberry, J. (1996). Improving representation of linguistic minorities in health surveys: A preliminary test of multilingual audio-CASI. *Public Health Reports, 111*, 276–279.

Zenilman, J. M., Weisman, C. S., Rompalo, A. M., Ellish, N., Upchurch, D. M., Hook, E. W., & Celentano, D. (1995). Condom use to prevent STDs: The validity of self-reported condom use. *Sexually Transmitted Diseases, 22*, 15–21.