

Confirmation of Family Cancer History Reported in a Population-Based Survey

Phuong L. Mai, Anne O. Garceau, Barry I. Graubard, Marsha Dunn, Timothy S. McNeel, Lou Gonsalves, Mitchell H. Gail, Mark H. Greene, Gordon B. Willis, Louise Wideroff

Manuscript received August 10, 2010; revised March 7, 2011; accepted March 9, 2011.

Correspondence to: Phuong L. Mai, MD, MS, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, 6120 Executive Blvd, EPS/7022, Rockville Pike, Bethesda, MD 20852 (e-mail: maip@mail.nih.gov).

Background Knowledge of family cancer history is essential for estimating an individual's cancer risk and making clinical recommendations regarding screening and referral to a specialty cancer genetics clinic. However, it is not clear if reported family cancer history is sufficiently accurate for this purpose.

Methods In the population-based 2001 Connecticut Family Health Study, 1019 participants reported on 20 578 first-degree relatives (FDR) and second-degree relatives (SDR). Of those, 2605 relatives were sampled for confirmation of cancer reports on breast, colorectal, prostate, and lung cancer. Confirmation sources included state cancer registries, Medicare databases, the National Death Index, death certificates, and health-care facility records. Sensitivity, specificity, positive predictive value, and negative predictive value were calculated for reports on lung, colorectal, breast, and prostate cancer and after stratification by sex, age, education, and degree of relatedness and used to estimate report accuracy. Pairwise *t* tests were used to evaluate differences between the two strata in each stratified analysis. All statistical tests were two-sided.

Results Overall, sensitivity and positive predictive value were low to moderate and varied by cancer type: 60.2% and 40.0%, respectively, for lung cancer reports, 27.3% and 53.5% for colorectal cancer reports, 61.1% and 61.3% for breast cancer reports, and 32.0% and 53.4% for prostate cancer reports. Specificity and negative predictive value were more than 95% for all four cancer types. Cancer history reports on FDR were more accurate than reports on SDR, with reports on FDR having statistically significantly higher sensitivity for prostate cancer than reports on SDR (58.9% vs 21.5%, $P = .002$) and higher positive predictive value for lung (78.1% vs 31.7%, $P < .001$), colorectal (85.8% vs 43.5%, $P = .004$), and breast cancer (79.9% vs 53.6%, $P = .02$).

Conclusions General population reports on family history for the four major adult cancers were not highly accurate. Efforts to improve accuracy are needed in primary care and other health-care settings in which family history is collected to ensure appropriate risk assessment and clinical care recommendations.

J Natl Cancer Inst 2011;103:788–797

Family health history is an important risk factor for several malignancies, cardiovascular disease, diabetes, and stroke (1) and reflects complex interactions among inherited genetic susceptibilities and shared environmental and behavioral factors. A comprehensive and accurate family cancer history is essential for cancer risk assessment in clinical and population surveillance settings (2,3). In the United States, cancer is a common disease; thus, the ability to accurately estimate risk for the more prevalent cancers, such as breast, prostate, and colorectal, and to devise appropriate screening and prevention plans, is an important aspect of public health practice (4,5).

Family members are not always aware of their relatives' health history. Cancer history might not be shared or discussed among family members for various reasons, ranging from geographical distance to a desire to protect family members from anxiety (6–9). Furthermore, even when cancer history is shared among family

members, the information disseminated is not always accurate, possibly because of the complexity of the diagnoses. Previous studies have shown that the accuracy of reported family cancer history is affected by many factors, including type of cancer, degree of relatedness, education, and sex (10).

In 2004, the US Surgeon General's Family Health History Initiative was launched to promote awareness and improve family history information ascertainment (11). However, family history is not accurately or consistently collected during primary care clinic visits (12–16). Family history, including cancer history, has an important role in clinical practice and improving health, but evidence regarding effective collection and utility is lacking (17). In addition, assessing the prevalence of family cancer history is important in estimating the demands for cancer screening and prevention activities on the US health service system. Furthermore, collection of accurate family history in population-based risk factor

surveillance surveys is critical in making valid estimates of the effect size of a positive family history on risk. In this report, we examine the accuracy of family history of common adult cancers in first-degree relatives (FDR) and second-degree relatives (SDR) as reported by the respondents in a population-based survey.

Methods

Cancer History Reports

The Family Health Study (FHS) was a representative random-digit-dial survey conducted in 2001 in the state of Connecticut. The purpose of the FHS was to develop a family cancer history questionnaire for surveillance purposes, to administer it to a sample of the general population, and to evaluate the accuracy of reports of cancer history in FDR and SDR by study respondents, by comparing these reports with cancer registry data and other health records. The state of Connecticut was chosen because it has the oldest population-based cancer registry in the United States, with records dating back to 1935, thus facilitating the process of confirming cancer reports. The study was approved by the Institutional Review Boards at the National Cancer Institute, Westat, Inc (the contract company that conducted the fieldwork), and the various state cancer registries that responded to our request for records.

Study details have been described elsewhere (18). In brief, a random sample of 11 982 telephone numbers in Connecticut was selected using list-assisted random-digit-dial telephone sampling (19,20). Attempts were made to obtain addresses for each selected telephone number. If an address was available, an introductory letter and a study pamphlet were sent before telephone contact. A total of 2418 residential households with at least one adult aged 25–64 years were identified; if more than one adult was identified, the individual with the most recent birthday was further screened for eligibility. Other eligibility criteria included being raised by at least one biological relative (to maximize the chance that a respondent would know about at least one lineage of the biological family) and having parents, or at least one parent and sibling, born or raised in the United States (to exclude respondents whose family cancer history could not be confirmed using US data sources).

Respondent Interviews. Two telephone interviews were conducted with the respondents. The first interview, lasting 20 minutes, on average, took place immediately following the screening questions and verbal consent (unless respondents requested otherwise). Respondents were asked for their full date of birth, personal cancer history (yes, no, or don't know) and if applicable, the type of cancer or, if unknown, where in the body it first occurred, and the age or year of diagnosis. Up to three primary cancers were obtained. Respondents were then asked to enumerate all biological FDR (parents, siblings, and children) and SDR (grandparents, uncles, aunts, nieces, and nephews), with the exception of grandchildren (because of the low cancer prevalence expected in young individuals). For each relative, first name, vital status, year of birth or age if living, year or age at death if deceased, cancer history, and if applicable, cancer type for up to three primary cancers, and year or age of diagnosis, were ascertained. After excluding subjects who were ineligible ($n = 304$), who declined participation

CONTEXTS AND CAVEATS

Prior knowledge

Family medical history is an important resource for health professionals in risk assessment, prevention, and treatment of cancer, but evidence on the accuracy and utility of reports of cancer by family members is lacking.

Study design

In a population-based survey, 1019 participants in the 2001 Connecticut Family Health Study reported on family history of breast, colorectal, prostate, and lung cancers in 20 578 first- and second-degree relatives. State cancer registries, Medicare databases, the National Death Index, death certificates, and health-care facility records were used to confirm these reports for a sample of 2605 relatives.

Contribution

Family medical histories were more accurate for first-degree than for second-degree relatives. Overall, reports of no history of cancer were highly accurate, but the accuracy of reported cancer diagnoses in relatives was low to moderate and varied by cancer type, with the highest accuracy for breast cancer and the lowest for colorectal cancer.

Implications

More accurate family medical histories might be available to health professionals once electronic health records and linked databases are in general use in the United States.

Limitations

Pathology report confirmation of cancer registry data was not available for all relatives. The study population may not be representative of the general US population because minorities and individuals of lower socioeconomic status were underrepresented and the age range was restricted to 25–64 years.

From the Editors

($n = 504$), and who could not be contacted ($n = 180$) or could not consent because of communication impediments ($n = 50$), a total of 1380 eligible respondents completed the first telephone interview, for a Council of American Survey Research Organizations response rate of 70%.

The second interview was carried out to obtain respondents' consent for investigators to contact sampled living relatives or proxies for deceased relatives and to gather respondent demographic data and personal identifiers. This interview was conducted within 1 month after the first interview, to allow time for random sampling of a subset of relatives whose cancer histories would be confirmed, because limited resources precluded confirmation of all relatives. The sampling process, in which up to six relatives were selected per family, has been described elsewhere (21) and is outlined in Table 1. Before being called, the respondents received a package containing sampled relatives' first names and reported cancer histories, key questions to be asked in the second interview, and a 60-minute calling card for optional long distance calls to sampled relatives regarding the study.

The length of the second interview varied according to the number of sampled relatives (mean = three relatives). Respondents were asked more detailed information about each sampled relative

Table 1. Categorization, definition, and criteria for selecting relatives to be included in the validation study*

Categories of relatives	Definition	Selection criteria
High-incidence FDR	Male FDR aged ≥ 55 y and female FDR aged ≥ 50 y	≤ 2 : sample one per family; >2 : sample two per family
High-incidence SDR	Male SDR aged ≥ 55 y and female SDR aged ≥ 50 y	≤ 10 : sample one per family; >10 : sample two per family
Low incidence cancer report relatives	Male relatives aged < 55 y and female relatives aged < 50 y reported to have cancer	If any: sample one per family
Low incidence noncancer report relatives	Male relatives aged < 55 y and female relatives aged < 50 y not reported to have cancer	Randomly selected from all respondent families, to a total number three times the total number of sampled low incidence cancer report relatives

* Relatives could potentially have more than one chance of being selected (ie, could be a relative of more than one respondent). This possibility was not adjusted for by weighting, so that these multiple-roster relatives were effectively overrepresented. FDR = first-degree relative; SDR = second-degree relative.

(full name and aliases, full date of birth, current or last known address, former states of residence for at least a year, and if deceased, month, year, state of residence, and marital status at time of death). For relatives with positive cancer reports, the facility where the diagnosis was made and state of residence at the time of diagnosis were collected. Permission to contact living relatives, or proxies of those deceased, was also requested. If no proxy was indicated, respondents were asked to provide the deceased relative's social security number (SSN). They were also asked to sign and return a medical records release form, subsequently sent by mail, if a reported cancer was diagnosed outside the Connecticut Tumor Registry (CTR) catchment area within the previous 7 years. Last, respondents' demographic information, including race, education, and income level, was collected. A total of 1019 respondents completed the second interview, for a Council of American Survey Research Organizations response rate of 74% (1019/1380 respondents).

Cancer History Confirmation

The 1019 respondents who completed both interviews reported a total of 20 578 relatives. Of the 2804 relatives selected for the cancer confirmation portion of this study, 10 requested to be removed from the study, 188 did not have full name available, and one lacked a living relative or proxy interview and had lived exclusively in an area where tumor registry data could not be obtained. Thus, cancer confirmation was attempted for a total of 2605 sampled relatives. Positive and negative cancer histories of these relatives were confirmed using five types of medical records systems: state tumor registries, Medicare claims databases, National Death Index (NDI), death certificates, and health-care facility records (Figure 1). All cancer codes were converted to *ICD-9-CM* (World Health Organization, *International Classification of Diseases, Ninth Revision, Clinical Modification*) codes for the analysis.

Generally, the critical personal identifiers for matching relatives included full name, date of birth, SSN, sex, full address, marital status, and if deceased, date or age at death. If SSN, the "gold standard" identifier, was missing, most confirmation sources offered alternative matching algorithms. Multiple data systems were searched per relative wherever inclusion criteria were met. Procedures for searching cancer diagnoses varied across systems, as described below. All personal identifiers were deleted before

analysis, thus eliminating the ability to repeat the matching process or identify respondents or relatives. If a relative's cancer reported by the respondent was confirmed by any of the sources, it was considered a true positive.

State Tumor Registries. Fifty-one state and regional tumor registries were contacted. Of these, 26 found matches, 20 found no matches, and five were unable to address the data request within the study time frame. The 26 registries found matches for 337 relatives out of 2586 submitted requests (Figure 1). Following local institutional review board approval, completion of confidentiality agreements and matching, registries returned de-identified data for both definitive and possible matches, which were linked back to the study database using a unique identification. Different methods of matching were used by state registries, ranging from manual review to sophisticated linkage software. Only definitive matches returned by the registries were used for analysis.

Medicare Claims Databases. The Centers for Medicare and Medicaid Services processes administrative claims to reimburse covered services provided to Medicare beneficiaries, most of whom are aged 65 years or older. Sampled relatives' information was submitted to the Centers for Medicare and Medicaid Services for electronic data linkage to claims data for the years 1984–2001 if they were living and older than 55 or if they died after 1975 at age 55 or more years (or age unknown) (Figure 1). Cutoffs of age 55 years instead of 65 and 1975 instead of 1984 were selected to accommodate age and date reporting errors. Relatives' identifiers were electronically linked to four national claims databases including Medical Provider Analysis and Review (inpatient hospitalizations), Hospice, Outpatient (hospital services), and Carrier (physician visits). Matching was determined using the Surveillance, Epidemiology, and End Results–Medicare match programs to create a finder file, which was then matched with a Health Insurance Claim Number that is unique to each Medicare beneficiary. Only de-identified data for definitive matches were returned and linked back to the study database using the unique study identification for each relative. The Medicare data use a combination of *ICD-9-CM* diagnosis codes, Healthcare Common Procedure Coding System codes, and revenue codes for each billed service. Each relative's Medicare claims were reviewed electronically for

cancer-related diagnoses or procedures codes. These records were flagged and reviewed manually by the study team. A total of 979 matches were found of 1974 requests submitted to the Medicare claims databases (Figure 1).

NDI Plus. The NDI Plus database, which provides electronic death certificate information on the US population from 1979 onwards, was searched to obtain primary and contributing causes of death. Personal identifiers were submitted for relatives who died after 1974 (to accommodate 5-year reporting error in year of death) through 2000 or whose vital status was unknown (Figure 1). The NDI Plus returned names and other identifiers of relatives or possible matches, along with the corresponding ICD codes indicating the cause of death. A final level of matching for each returned record was assigned by the investigators based on how well the personal identifiers matched with the relative's information. A returned record was determined to be a definitive match for a relative if any of the following criteria were met: 1) all of its identifiers matched those of the relative exactly; 2) the SSN matched exactly and at least four of the following seven key linkage variables matched with those of the relative: first and last name, month and year of death, month and year of birth, and state of residence; 3) it matched with any of the nine digits of the SSN and all seven key linkage variables; or 4) the SSN was not available and it matched with all seven key linkage variables. Only definitive matches (444 out of 919 submitted) were included in the analyses (Figure 1).

Death Certificates. Death certificates were sought for deceased sampled relatives who did not meet the inclusion criteria for the CTR or the NDI Plus, did not match in the NDI Plus, or had died

in New York City where NDI Plus cannot release cause of death (Figure 1). Individual state and New York City vital statistics offices were contacted to request the death certificates, and available personal identifiers were submitted. Of the 58 requests submitted, 38 matches were received from 12 states and New York City (Figure 1).

Medical Records. If living relatives, proxies, or respondents had returned a medical records release form, physician offices, hospitals, or other facilities were contacted to confirm cancer diagnoses. Out of 112 medical records requested, a total of 44 were received from providers (Figure 1). Forms were only requested for cancers diagnosed outside the CTR catchment area within the previous 7 years because low yield was anticipated as a result of facility closures, finite record storage, and work burden. Medical records were not pursued for those with a negative cancer history.

Living Relative/Proxy Interviews. Permission to contact relatives or their proxies for the living relative/proxy interview and location information were obtained for 1481 of the sampled relatives (Figure 1). Each of these sampled relatives or proxies was sent an invitational letter, which included a toll-free number that they could call to opt out of the upcoming interview or the entire study, study pamphlet, and a \$10 advance compensation. Relatives/proxies were contacted within approximately 1 week by telephone for a 10-minute interview. The purpose of the interview was to obtain verbal consent for participation in the study, personal identifiers for the confirmation process, and self-reported cancer histories. Of these 1481 relatives, 203 were not locatable, 160 declined to complete the interview, and 1118 (640 living relatives

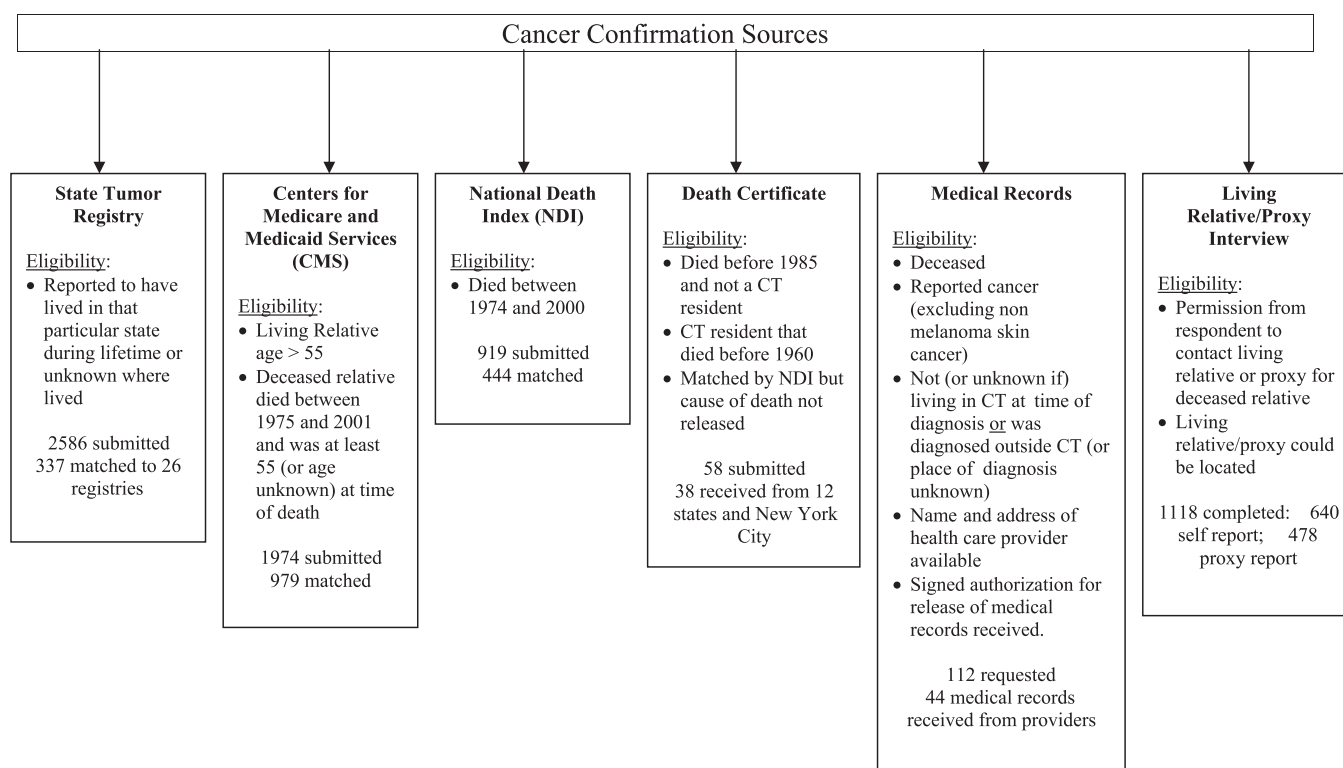


Figure 1. Data sources, with inclusion criteria, used to confirm relatives' cancer diagnoses as reported by respondents. The numbers of records submitted and matched are shown for each source. CT = Connecticut.

and 478 proxies) completed the interview (Figure 1). Self-reported cancer diagnosis information obtained during the relative/proxy interview was used as one of the cancer confirmation sources.

Statistical Analysis

The primary objective of the study was to estimate the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of reports on family cancer history by study participants. Each sampled relative was assigned a sampling weight that incorporates the sampling rates of the respondent and the relative to adjust for differential selection probabilities. A replicate weight approach based on the delete-one jackknife method was used in standard error estimation (22). The delete-one jackknife approach was used for the estimation of standard errors, which was then applied in the computation of confidence intervals and of *P* values for statistical comparisons, to take into account of the fact that reports of cancer history of multiple relatives from the same respondent were nonindependent.

Sensitivity, specificity, PPV, and NPV (Table 2) and the corresponding 95% confidence intervals (CI) were calculated for reports on lung, colorectal, breast, and prostate cancers. All analyses were weighted by the sample weights. The same measures of accuracy were also calculated after stratification by selected respondent and relative characteristics. Respondent characteristics included sex, age group (25–44 vs 45–64 years), and education (high school, vocational, or technical school graduate, or less vs at least some college). Relative characteristics include degree of relatedness to respondent (FDR vs SDR) and whether a living relative/proxy interview was completed. Two-sided pairwise *t* tests were used to evaluate for differences between the two strata in each stratified analysis. SAS version 9.1 (SAS Institute, Inc, NC; 2004) and SAS-callable SUDAAN version 10.0.1 (Research Triangle Institute, NC, 2009) statistical software were used to conduct the analyses and estimate the standard errors. All reported *P* values are two-sided.

Results

Cancer Confirmation

SSNs were available for 42% of the 2605 sampled relatives included in the analyses, including 52% of living and 31% of deceased relatives. Other personal identifiers obtained were birth month (86% of living and 69% of deceased), address (83% of living and 67% of deceased), marital status at time of death (99% of deceased), month and year of death (72% and 88% of deceased, respectively), and age at death (41% of deceased). Year of birth and sex were available for all 2605 relatives.

Lung Cancer Reports

Fifty-five of the 105 reported lung cancers were confirmed. The overall sensitivity was 60.2% (95% CI = 46.5 to 72.4), specificity was 98.3% (95% CI = 97.3 to 98.9), PPV was 40.0% (95% CI = 26.4 to 55.4), and NPV was 99.2% (95% CI = 98.8 to 99.5) (Table 3). Reports on FDR had a statistically significantly higher specificity (99.7%, 95% CI = 99.3 to 99.9, *P* < .001) and PPV (78.1%, 95% CI = 59.3 to 89.7, *P* < .001) than reports on SDR (specificity: 97.5%, 95% CI = 96.1 to 98.5; PPV: 31.7%, 95% CI = 18.0 to 49.6). Reporting sensitivity was also higher for FDR than for SDR, but the difference was not statistically significant. Compared with reports on relatives for whom a living relative/proxy interview was not completed, reports on relatives for whom the interview was completed had a non-statistically significantly higher PPV (63.2% vs 25.7%, *P* = .08) and a statistically significantly lower NPV (98.6% vs 99.7%, *P* = .02). When examined by respondents' characteristics, reports by female respondents had a PPV of 54.0% (95% CI = 35.3 to 71.6), whereas that of reports by male respondents was 29.2% (95% CI = 13.6 to 52.0), a non-statistically significant difference. Younger age of respondents was associated with a higher sensitivity (73.2% vs 34.9%, *P* < .001). Respondents' level of education did not affect the sensitivity, specificity, PPV, or NPV.

A false-negative report for lung cancer was documented for 28 relatives. Of these, 14 were not reported by the respondents to have had a cancer at all; six were reported to have had other types of cancer, which were confirmed; and eight were reported to have had a cancer at an incorrect site, of which two might have been metastases (one brain and one bone cancer), and three were sites that could be confused with lung cancer, given their locations (throat, esophageal, and neck).

Colorectal Cancer Reports

Of the 58 colorectal cancers reported, 40 were confirmed. The sensitivity was 27.3% (95% CI = 16.7 to 41.3), specificity was 99.4% (95% CI = 98.8 to 99.7), PPV was 53.5% (95% CI = 33.0 to 72.8), and NPV was 98.1 (95% CI = 96.6 to 98.9) (Table 3). Reports on FDR had a statistically significantly higher PPV than reports on SDR (85.8% vs 43.5%, *P* = .004). The specificity for reports on FDR was also higher (99.9% vs 99.1%, *P* = .02). The PPV for reports on relatives for whom a living relative/proxy interview was available was statistically significantly higher than that on relatives for whom the interview was not completed (90.4% vs 20.0%, *P* < .001). Although male sex and younger age of respondents had a higher PPV, none of the respondents' characteristics analyzed was statistically significantly associated with any of the measures of accuracy.

Table 2. Definitions of measures of report accuracy*

Identified in any of the confirmation sources	Report of cancer by respondent		
	Present	Absent	
Yes	TP	FN	Sensitivity = TP/TP + FN
No	FP	TN	Specificity = TN/TN + FP
	PPV = TP/TP + FP	NPV = TN/TN + FN	

* FN = false negative; FP = false positive; NPV = negative predictive value; PPV = positive predictive value; TN = true negative; TP = true positive.

Table 3. Sensitivity, specificity, PPV, and NPV of lung, colorectal, breast, and prostate cancer reports*

Categories of relatives (total N, No. of positive reports, No. of true-positive reports)†	% (95% CI)			
	Sensitivity	Specificity	PPV	NPV
Lung cancer				
All eligible relatives (2605, 105, 55)	60.2 (46.5 to 72.4)	98.3 (97.3 to 98.9)	40.0 (26.4 to 55.4)	99.2 (98.8 to 99.5)
First-degree relatives (1361, 47, 36)	71.3 (52.8 to 84.7)	99.7 (99.3 to 99.9)	78.1 (69.3 to 89.7)	99.6 (99.1 to 99.8)
Second-degree relatives (1244, 58, 19)	55.5 (37.3 to 72.3)	97.5 (96.1 to 98.5)	31.7 (18.0 to 49.6)	99.1 (98.5 to 99.4)
		<i>P</i> < .001‡	<i>P</i> < .001‡	
Completed living relative/proxy interview (1118, 38, 31)	55.7 (38.3 to 71.9)	99.0 (96.3 to 99.7)	63.2 (24.0 to 90.4)	98.6 (97.6 to 99.2)
Not completed living relative/proxy interview (1487, 67, 24)	68.5 (41.6 to 86.9)	97.9 (96.6 to 98.7)	25.7 (13.1 to 44.2)	99.7 (99.3 to 99.8)
				<i>P</i> = .02§
Relatives of male respondents (1004, 44, 20)	55.0 (31.2 to 76.7)	97.6 (95.3 to 98.8)	29.2 (13.6 to 52.0)	99.2 (98.4 to 99.6)
Relatives of female respondents (1601, 61, 35)	64.4 (45.3 to 79.8)	98.9 (98.1 to 99.4)	54.0 (35.3 to 71.6)	99.3 (98.8 to 99.6)
Relatives of respondents aged 25–44 y (1017, 53, 30)	73.2 (56.3 to 85.3)	97.9 (96.0 to 98.9)	48.3 (27.8 to 69.3)	99.3 (98.5 to 99.6)
Relatives of respondents aged 45–64 y (1588, 52, 25)	34.9 (20.9 to 52.3)	98.7 (97.4 to 99.3)	23.7 (12.0 to 41.5)	99.2 (98.6 to 99.6)
	<i>P</i> < .001			
Relatives of respondents with high school/vocational/technical school degree or less (799, 37, 19)	64.2 (38.2 to 83.8)	97.9 (95.0 to 99.1)	39.3 (15.2 to 69.9)	99.2 (98.5 to 99.6)
Relatives of respondents with some college education or more (1804, 68, 36)	58.1 (40.7 to 73.7)	98.5 (97.4 to 99.1)	40.5 (24.2 to 59.2)	99.3 (98.7 to 99.6)
Colorectal cancer				
All eligible relatives (2605, 58, 40)	27.3 (16.7 to 41.3)	99.4 (98.8 to 99.7)	53.5 (33.0 to 72.8)	98.1 (96.6 to 98.9)
First-degree relatives (1361, 34, 29)	26.8 (6.2 to 66.9)	99.9 (99.6 to 100.0)	85.8 (63.9 to 95.3)	97.9 (93.8 to 99.3)
Second-degree relatives (1244, 24, 11)	27.6 (15.1 to 45.1)	99.1 (98.1 to 99.5)	43.5 (21.6 to 68.3)	98.2 (97.1 to 98.8)
		<i>P</i> = .02‡	<i>P</i> = .004‡	
Completed living relative/proxy interview (1118, 26, 25)	34.9 (14.6 to 62.8)	99.8 (98.8 to 100.0)	90.4 (49.3 to 98.9)	97.3 (93.1 to 98.9)
Not completed living relative/proxy interview (1487, 32, 15)	14.4 (7.3 to 26.3)	99.0 (98.1 to 99.5)	20.0 (8.1 to 41.4)	98.6 (97.6 to 99.2)
		<i>P</i> = .04§	<i>P</i> < .001§	
Relatives of male respondents (1004, 26, 19)	27.5 (11.5 to 52.4)	99.4 (98.4 to 99.8)	60.1 (30.4 to 83.8)	97.6 (94.1 to 99.0)
Relatives of female respondents (1601, 32, 21)	27.1 (13.9 to 46.0)	99.3 (98.3 to 99.7)	46.0 (17.7 to 77.2)	98.5 (97.6 to 99.0)
Relatives of respondents aged 25–44 y (1017, 22, 15)	29.0 (13.6 to 51.4)	99.4 (98.5 to 99.8)	63.1 (33.2 to 85.5)	97.5 (94.7 to 98.8)
Relatives of respondents aged 45–64 y (1588, 36, 25)	24.7 (14.6 to 38.8)	99.3 (98.4 to 99.7)	42.0 (16.7 to 72.4)	98.5 (97.5 to 99.1)
Relatives of respondents with high school/vocational/technical school degree or less (799, 14, 9)	30.3 (12.5 to 57.1)	99.4 (97.2 to 99.9)	52.8 (9.1 to 92.6)	98.4 (96.9 to 99.1)
Relatives of respondents with some college education or more (1804, 44, 31)	26.2 (13.6 to 44.3)	99.4 (98.7 to 99.7)	53.7 (30.8 to 75.1)	97.9 (95.8 to 99.0)
Breast cancer				
All eligible relatives (1433, 119, 77)	61.1 (48.3 to 72.5)	98.1 (97.0 to 98.8)	61.3 (47.1 to 73.8)	98.1 (97.1 to 98.7)
First-degree relatives (750, 60, 45)	64.9 (49.6 to 77.6)	99.0 (98.2 to 99.5)	79.9 (66.1 to 89.0)	97.9 (96.4 to 98.7)
Second-degree relatives (683, 59, 32)	59.0 (39.9 to 75.7)	97.7 (96.1 to 98.7)	53.6 (35.0 to 71.3)	98.1 (96.7 to 98.9)
		<i>P</i> = .02‡	<i>P</i> = .02‡	
Completed living relative/proxy interview (640, 66, 59)	66.4 (51.6 to 78.6)	99.1 (97.9 to 99.7)	86.8 (75.2 to 95.4)	96.6 (94.5 to 98.0)
Not completed living relative/proxy interview (793, 53, 18)	49.3 (26.7 to 72.2)	97.5 (95.7 to 98.6)	31.9 (16.8 to 52.0)	98.8 (97.7 to 99.4)
		<i>P</i> = .02§	<i>P</i> < .001§	<i>P</i> = .05§
Relatives of male respondents (538, 44, 28)	56.6 (35.5 to 75.5)	98.0 (95.6 to 99.1)	57.6 (32.8 to 79.0)	97.9 (95.9 to 98.9)
Relatives of female respondents (895, 75, 49)	64.5 (48.6 to 77.8)	98.1 (96.7 to 99.0)	64.1 (46.5 to 78.5)	98.2 (97.0 to 98.9)
Relatives of respondents aged 25–44 y (546, 45, 31)	68.9 (51.8 to 82.0)	98.2 (96.1 to 99.2)	67.8 (43.4 to 85.3)	98.3 (96.9 to 99.1)
Relatives of respondents aged 45–64 y (887, 74, 46)	52.6 (35.2 to 69.4)	97.9 (96.4 to 98.8)	53.9 (36.6 to 70.4)	97.8 (96.4 to 98.7)
Relatives of respondents with high school/vocational/technical school degree or less (435, 36, 21)	71.8 (41.9 to 90.9)	97.7 (95.1 to 99.0)	56.4 (32.4 to 77.8)	98.8 (97.1 to 99.5)
Relatives of respondents with some college education or more (997, 83, 56)	57.7 (42.4 to 71.6)	98.2 (96.8 to 99.0)	63.5 (44.2 to 79.2)	97.7 (96.4 to 98.6)

(Table continues)

Table 3 (Continued).

Categories of relatives (total N, No. of positive reports, No. of true-positive reports)†	% (95% CI)			
	Sensitivity	Specificity	PPV	NPV
Prostate cancer				
All eligible relatives (1172, 57, 42)	32.0 (21.4 to 45.0)	98.6 (97.2 to 99.3)	53.4 (35.7 to 70.3)	96.9 (94.8 to 97.8)
First-degree relatives (611, 44, 36)	58.9 (43.1 to 73.1)	98.9 (96.2 to 99.7)	64.7 (30.7 to 88.4)	98.6 (96.9 to 99.4)
Second-degree relatives (561, 13, 6)	21.5 (9.5 to 41.7)	98.3 (96.3 to 99.3)	45.0 (18.1 to 75.2)	95.1 (92.6 to 96.8)
Completed living relative/proxy interview (478, 31, 29)	41.1 (25.7 to 58.5)	99.2 (96.5 to 99.8)	80.0 (44.3 to 95.3)	95.7 (91.1 to 98.0)
Not completed living relative/proxy interview (694, 26, 13)	14.5 (7.2 to 26.9)	98.1 (96.0 to 99.1)	19.0 (8.5 to 37.1)	97.4 (95.4 to 98.5)
Relatives of male respondents (466, 26, 19)	28.6 (14.2 to 49.3)	98.5 (95.7 to 99.5)	54.4 (25.3 to 80.8)	95.7 (91.8 to 97.8)
Relatives of female respondents (706, 31, 23)	37.1 (23.1 to 53.7)	98.7 (96.5 to 99.5)	52.3 (24.9 to 78.4)	97.5 (95.9 to 98.5)
Relatives of respondents aged 25–44 y (471, 22, 13)	31.2 (16.0 to 51.9)	98.4 (96.0 to 99.4)	56.4 (28.1 to 81.1)	95.7 (93.1 to 97.4)
Relatives of respondents aged 45–64 y (701, 35, 29)	33.0 (18.1 to 52.4)	98.7 (96.3 to 99.5)	50.4 (26.5 to 74.2)	97.3 (94.6 to 98.7)
Relatives of respondents with high school/vocational/technical school degree or less (364, 13, 9)	33.0 (9.8 to 69.2)	98.8 (95.4 to 99.7)	46.7 (10.1 to 87.3)	97.8 (94.6 to 99.1)
Relatives of respondents with some college education or more (807, 44, 33)	31.7 (20.5 to 45.6)	98.5 (96.6 to 99.4)	55.6 (34.6 to 74.9)	96.1 (93.5 to 97.6)

* Weighted for the population of the state of Connecticut. CI = confidence interval; NPV = negative predictive value; PPV = positive predictive value.

† Unweighted numbers.

‡ Two-sided pairwise *t* tests between reports on first-degree relatives vs second-degree relatives.

§ Two-sided pairwise *t* tests between completed living relative/proxy interviews vs not completed living relative/proxy interviews.

|| Two-sided pairwise *t* tests of reports by respondents aged 25–44 years vs respondents aged 45–64 years.

Discussion

We assessed the accuracy of reported family history of breast, colorectal, prostate, and lung cancers. Overall, specificity and NPV were high, whereas sensitivity and PPV were low to moderate.

Prostate Cancer Reports

There were 1172 men in the confirmation study. Forty-two of the 57 reported prostate cancers were confirmed, resulting in a sensitivity of 32.0% (95% CI = 21.4 to 45.0), a specificity of 98.6% (95% CI = 97.2 to 99.3), a PPV of 53.4% (95% CI = 35.7 to 70.3), and an NPV of 96.9% (95% CI = 94.8 to 97.8) (Table 3). Compared with reports on SDR, reports on FDR had a higher sensitivity (58.9% vs 21.5%, $P = .002$) and a higher NPV (98.6% vs 95.1%, $P = .01$). Prostate cancer reports on relatives with a living relative/proxy interview had a higher sensitivity (41.1% vs 14.5%, $P = .01$) and a considerably higher PPV (80.0% vs 19.0%, $P < .001$). Similar to reports on breast cancer, none of the measures of accuracy for prostate cancer varied statistically significantly by the respondents' characteristics examined.

Thirty-four of the 44 relatives with a false-negative report were reported not to have had cancer. Of the 10 relatives reported to have had a cancer, seven were accurately reported for a cancer other than prostate. The other three false-negative reports were stomach cancer, abdominal cancer, and cancer of unknown origin. No differences in any of the report accuracy measures for any of these cancers were observed when examined by respondents' personal history of cancer (data not shown).

Breast Cancer Reports

History of breast cancer was collected on 1433 female relatives (Table 3). A total of 119 breast cancers were reported, of which 77 were confirmed, yielding a sensitivity of 61.1% (95% CI = 48.3 to 72.5), a specificity of 98.1% (95% CI = 97.0 to 98.8), a PPV of 61.3% (95% CI = 47.1 to 73.8), and an NPV of 98.1% (95% CI = 97.1 to 98.7). Similar to reports on lung and colorectal cancers, breast cancer reports on FDR had a statistically significantly higher specificity (99.0% vs 97.7%, $P = .02$) and PPV (79.9% vs 53.6%, $P = .02$). Reports on relatives with a living relative/proxy interview also had a higher PPV (88.8% vs 31.9%, $P < .001$). The respondents' characteristics examined did not affect any of the measures of accuracy.

No cancer was reported for 28 of the 42 false-negative reports for breast cancer. Seven relatives had other types of cancer reported accurately, but no breast cancer was mentioned. Of the seven false-negative reports with an inaccurate site, one was reported as a benign breast lump and one may have represented a metastatic site (liver cancer).

Of the 51 false-negative reports for colorectal cancer, 25 were reported as not having had any cancer at all. Of the cancers reported at a different site, 12 were accurately reported for other types of cancer (ie, these relatives had more than one cancer recorded in the confirmation sources), but a diagnosis of colorectal cancer was not mentioned. The other cancers reported were mostly either within the gastrointestinal tract or in the abdominal area. The two reports of lung cancer and one of bone cancer could represent the site of metastasis. Two reports were cancer of unknown type.

One or more of the measures of accuracy was affected by cancer type and by degree of relatedness to the respondents. Breast cancer reports had the highest sensitivity and PPV, whereas colorectal cancer had the lowest sensitivity and PPV. Except for prostate cancer, PPV was statistically significantly higher for reports on FDR than for those on SDR. In general, cancer history obtained on relatives with a completed living relative/proxy interview had higher measures of accuracy.

The sensitivity of family history reports for the four adult cancers in our study was lower than previously reported by other studies, particularly for colorectal and prostate cancers. In one review (23), sensitivity of family history reports ranged from 72% to 95% for breast cancer, 30% to 90% for colorectal cancer, and 47% to 79% for prostate cancer (23). The lower sensitivity observed in the Connecticut FHS could be because of its population-based design. Many previous estimates of family cancer history reporting accuracy used information collected in specialty clinics or from respondents who themselves had cancer, who were perhaps more knowledgeable about their family cancer history than the general population. No differences in reporting sensitivity by respondents' personal cancer history were observed in the Connecticut FHS (data not shown), but the number of cancer-affected respondents was small, and most of the cancers were non-melanoma skin cancer. We defined false negative as a report of no cancer for a relative found to have cancer by any of the confirmation sources; an accurate report of other cancers for relatives with more than one cancer diagnosis, but the cancer of interest was not mentioned; or a report of cancer at a different site. Overall, the numbers of reports of inaccurate cancer type were small, and some of the sites named were within the same system or anatomical region or could represent a metastasis. Prostate cancer had the most false-negative reports classified as "no cancer reports". Perhaps prostate cancer is less openly discussed in families, and men are generally less willing to discuss health issues or disseminate health information (9), leading to low awareness among relatives of this cancer diagnosis.

Overall, PPV was low, ranging from 40% (lung cancer) to 61.3% (breast cancer). Higher PPV has been reported previously, particularly for breast cancer (24). Potential reasons for false-positive reports include a misunderstanding of the benign vs malignant nature of the tumor (eg, fibroadenoma reported as breast cancer) and confusion between primary and metastatic site (eg, lung metastasis vs primary lung cancer). Because family history can affect the recommendations for breast and colorectal cancer screening (5), a low PPV could lead to overscreening. Furthermore, family history is essential in cancer risk estimation (25,26) and evaluation for hereditary cancer syndromes (27); consequently, inaccurate family history reports could lead to inappropriate risk management recommendations or unnecessary referral for genetic evaluation. Thus, there is a need to promote family cancer history awareness and to find better tools to capture it accurately, to ensure that appropriate risk assessment and clinical care recommendations can be made. One potential approach to improving family cancer history awareness, and consequently, family history report accuracy, is to raise general knowledge of cancers within the community. Improved knowledge about cancer might encourage people to be more willing to communicate about it with others, either when sharing information about their own diagnoses or when asking for information about their relatives' diagnoses.

Specificity and NPV were very high for all four cancers. Cancer diagnoses were rare among relatives in this study; thus, this finding is not surprising. Furthermore, our confirmation sources might have missed some cancer diagnoses. Consequently, some of the actual false-negative reports might have been misclassified as true negative because they were never confirmed.

We observed that accuracy varied by cancer type, with breast cancer reports being the most accurate, consistent with prior literature (24,28,29), which presumably reflects different levels of respondent awareness by cancer types. Women in general are more likely to be the family disseminator of health information (9,30) and thus might be more likely to share information about their own breast cancer diagnosis. Within families, some cancers, such as colorectal (14), might be less openly discussed, leading to decreased awareness and inaccurate reports.

Our results showed greater, although non-statistically significant, sensitivity and PPV for reported family history of breast, prostate, and lung cancers by female respondents. The association between respondents' sex and accuracy of reported family cancer history has been inconsistent, with some studies showing no differences (15,28) and others showing more accurate reports by women (31,32). We found that measures of accuracy were associated with degree of relatedness, as previously reported (24). Closely related relatives are more likely to share information. In addition, SDR include grandparents, who might be deceased, making their health information less readily available.

Cancer history information obtained on relatives for whom a living relative/proxy interview was also completed had higher measures of accuracy, which could be an artifact of verification method. Because the living relative/proxy interview served as a cancer confirmation source, positive reports on these relatives would be more likely to be confirmed. However, the ability to obtain consent for the interview might be an indication that the respondent kept close contact with the relative and thus was more knowledgeable about the relative's health history.

Our study has several strengths. First, this was a population-based survey in which we attempted to confirm all cancers reported on a randomly selected sample of relatives. The respondents in this study were more representative of the population seen in primary care clinics, where initial cancer risk assessment usually takes place, and of population control subjects randomly sampled into case-control studies. Therefore, our observation of low to moderate sensitivity and PPV for family cancer history may have important implications regarding the potential extent of reporting error in population surveillance, risk assessment, and risk estimation. Second, a large percentage of the sampled relatives reported living all their lives in Connecticut. As a result, we were able to effectively use the CTR as the main confirmation source. In addition, a small but important number lived in nearby states with long-standing tumor registries. Finally, this study focused on the four major adult cancers most common in the United States and whose risk has been shown to be affected by a positive family history.

There were also some limitations to this study. We were not able to obtain pathology report-based cancer registry data (the "gold standard" for assessing accuracy of cancer history reports) for all sampled relatives because of missing SSN or other critical personal identifiers or lack of a participating tumor registry where

the relative resided. In addition, there were variations in matching methods, percentage of pathology-confirmed cancers, and start date across state registries. The methods used in this study have been used as the standard in several prior studies and represent the best confirmation sources available (23). Except for cancer registries, all the other confirmation sources have limitations. When cancer is the direct cause of death, death certificates are accurate and can serve as an adequate source of cancer confirmation (33,34), but when cancer is unrelated to the immediate cause of death, a cancer history would not necessarily be recorded. For example, in the US Radiologic Technologists Study, death certificates did not identify 35.2% of cancer patients identified by cancer registries (35). Consequently, using death certificates as the only source of cancer confirmation can potentially result in an overestimation of false-positive reports (ie, falsely low PPV). Alternatively, data abstracted from Medicare claims databases are sensitive for cancer diagnoses (36,37) but might include other diagnoses or procedures coded as cancers (38), resulting in a falsely low sensitivity. Furthermore, self-reported cancers or cancers reported by proxies might not be accurate because one cancer can often be mistaken for another, depending on the location and nature of the disease. In this study, for the subset of relatives who lived only in Connecticut, there were cancers identified in at least one of the other sources, specifically by the Centers for Medicare and Medicaid Services, which were not recorded in the CTR (data not shown). It is unclear to what extent cancer diagnoses were over reported by administrative claims databases and might have contributed to the observed low sensitivity.

Alternatively, it is inevitable that some of the cancer reports were accurate but were not confirmed by any of the sources available and thus were classified as false positives. It is unclear how much of an impact this classification might have had on the results; however, we made every attempt to match the relatives with all of the confirmation sources available.

Although the study was population based, the participants were not strictly representative of Connecticut or the United States overall (18), with individuals of minority and low socioeconomic status being underrepresented. Thus, our results are most generalizable to the non-Hispanic white population of Connecticut and similar populations. In addition, because the eligibility criteria restricted the respondents to age 25–64, the results might not be generalizable outside of this age range. Furthermore, the interview data are approximately 10 years old; however, we are not aware of any data suggesting that accuracy of family cancer reports has markedly changed in recent years. Finally, although respondents were not asked to prepare for the first interview, some may have chosen to gather family history information anyway, thus increasing reporting accuracy and limiting extrapolation of our findings to primary care or research settings that lack opportunities for advance ascertainment.

In summary, the sensitivity and PPV of a reported family history of lung, colorectal, breast, and prostate cancers in this population-based survey were low to moderate, especially among SDR, but the specificity and NPV were high. Given that the population from which we sampled is similar to primary care populations, the results of this study suggest that family cancer history collected in the primary care setting might be useful as an initial screening tool,

and, if positive, confirmation of the reported cancers is needed for the purpose of making cancer screening recommendations or referral to a specialty clinic. Verification of family cancer history can be cumbersome, given the lack of a readily available electronically linkable source. Promoting awareness, encouraging people to ask questions about cancer in their family, and using pre-visit information collection tools such as that proposed in the Surgeon General's Family Health History Initiative might help improve family cancer history report accuracy.

References

1. Yoon PW, Scheuner MT, Jorgensen C, Khoury MJ. Developing Family Healthware, a family history screening tool to prevent common chronic diseases. *Prev Chronic Dis*. 2009;6(1):A33.
2. Yoon PW, Scheuner MT, Peterson-Oehlke KL, Gwinn M, Faucett A, Khoury MJ. Can family history be used as a tool for public health and preventive medicine? *Genet Med*. 2002;4(4):304–310.
3. Freedman AN, Graubard BI, Rao SR, McCaskill-Stevens W, Ballard-Barbash R, Gail MH. Estimates of the number of US women who could benefit from tamoxifen for breast cancer chemoprevention. *J Natl Cancer Inst*. 2003;95(7):526–532.
4. US Preventive Services Task Force. Guide to Clinical Preventive Services. <http://www.uspreventiveservicestaskforce.org/adultrec.htm#cancer>. Accessed December 22, 2009.
5. Smith RA, Cokkinides V, Brooks D, Saslow D, Brawley OW. Cancer screening in the United States, 2010: a review of current American Cancer Society guidelines and issues in cancer screening. *CA Cancer J Clin*. 2010;60(2):99–119.
6. Loescher LJ, Crist JD, Siaki LA. Perceived intrafamily melanoma risk communication. *Cancer Nurs*. 2009;32(3):203–210.
7. Matthews AK, Cummings S, Thompson S, Wohl V, List M, Olopade OI. Genetic testing of African Americans for susceptibility to inherited cancers: use of focus groups to determine factors contributing to participation. *J Psychosocial Oncol*. 2000;18(2):1–19.
8. Kenen R, Arden-Jones A, Eeles R. We are talking, but are they listening? Communication patterns in families with a history of breast/ovarian cancer (HBOC). *Psychooncology*. 2004;13(5):335–345.
9. Koehly LM, Peters JA, Kenen R, et al. Characteristics of health information gatherers, disseminators, and blockers within families at risk of hereditary cancer: implications for family health communication interventions. *Am J Public Health*. 2009;99(12):2203–2209.
10. Qureshi N, Wilson B, Santaguida P, et al. *Collection and Use of Cancer Family History in Primary Care*. Rockville, MD: Agency for Healthcare Research and Quality; 2007. Evidence Report/Technology Assessment No. 159 (prepared by the McMaster University Evidence-based Practice Center, under Contract No. 290-02-0020). AHRQ Publication No. 08-E001.
11. US Department of Health and Human Services. Surgeon General's Family Health History Initiative. <http://www.hhs.gov/familyhistory/>. Accessed December 22, 2009.
12. Murff HJ, Byrne D, Syngal S. Cancer risk assessment: quality and impact of the family history interview. *Am J Prev Med*. 2004;27(3):239–245.
13. Sifri RD, Wender R, Paynter N. Cancer risk assessment from family history: gaps in primary care practice. *J Fam Pract*. 2002;51(10):856. <http://www.jfponline.com/Pages.asp?AID=1965&issue=October%202002&UID=51:000>. Accessed June 12, 2010.
14. Mitchell RJ, Brewster D, Campbell H, et al. Accuracy of reporting of family history of colorectal cancer. *Gut*. 2004;53(2):291–295.
15. Sijmons RH, Boonstra AE, Reefhuis J, et al. Accuracy of family history of cancer: clinical genetic implications. *Eur J Hum Genet*. 2000;8(3):181–186.
16. Acheson LS, Wiesner GL, Zyzanski SJ, Goodwin MA, Stange KC. Family history-taking in community family practice: implications for genetic screening. *Genet Med*. 2000;2(3):180–185.
17. Berg AO, Baird MA, Botkin JR, et al. National Institutes of Health State-of-the-Science Conference Statement: Family History and Improving Health. *Ann Intern Med*. 2009;151(12):872–877.

18. Garceau A, Wideroff L, McNeel T, Dunn M, Graubard BI. Population estimates of extended family structure and size. *Community Genet.* 2008;11(6):331–342.
19. Levy PS, Lemeshow S. *Sampling of Populations: Methods and Applications*. 3rd ed. New York, NY: John Wiley and Sons, Inc; 1999.
20. Tucker C, Casady R, Lepkoski J. Sample allocation for stratified telephone sample designs. *Proceedings of the Survey Research Methods Section, ASA.* 1992:291–296.
21. Rizzo L, Graubard B, DiGaetano R, Wideroff L. Establishing rates for the family history of Cancer Validation Study. *Proceedings of the Survey Research Methods Section, ASA.* 2000:683–687.
22. Rust KF, Rao JN. Variance estimation for complex surveys using replication techniques. *Stat Methods Med Res.* 1996;5(3):283–310.
23. Qureshi N, Wilson B, Santaguida P, et al. *NIH State-of-the-Science Conference: Family History and Improving Health*. Rockville, MD: AHRQ, 2009. AHRQ Publication No. 09-E016. Evidence Report/Technology Assessment No. 186. (Prepared by the McMaster University Evidence-based Practice Center, under Contract No. 290-2007-10060-I).
24. Murff HJ, Spigel DR, Syngal S. Does this patient have a family history of cancer? An evidence-based analysis of the accuracy of family cancer history. *JAMA.* 2004;292(12):1480–1489.
25. Freedman AN, Slattey ML, Ballard-Barbash R, et al. Colorectal cancer risk prediction tool for white men and women without known susceptibility. *J Clin Oncol.* 2009;27(5):686–693.
26. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst.* 1989;81(24):1879–1886.
27. Robson ME, Storm CD, Weitzel J, Wollins DS, Offit K. American Society of Clinical Oncology policy statement update: genetic and genomic testing for cancer susceptibility. *J Clin Oncol.* 2010;28(5):893–901.
28. Ziogas A, Anton-Culver H. Validation of family history data in cancer family registries. *Am J Prev Med.* 2003;24(2):190–198.
29. Douglas FS, O'Dair LC, Robinson M, Evans DGR, Lynch SA. The accuracy of diagnoses as reported in families with cancer: a retrospective study. *J Med Genet.* 1999;36(4):309–312.
30. Hay J, Shuk E, Brady MS, Berwick M, Ostroff J, Halpern A. Family communication after melanoma diagnosis. *Arch Dermatol.* 2008;144(4):553–554.
31. Kerber RA, Slattey ML. Comparison of self-reported and database-linked family history of cancer data in a case-control study. *Am J Epidemiol.* 1997;146(3):244–248.
32. Glanz K, Grove J, Le Marchand L, Gotay C. Underreporting of family history of colon cancer: correlates and implications. *Cancer Epidemiol Biomarkers Prev.* 1999;8(7):635–639.
33. Ederer F, Geisser MS, Mongin SJ, Church TR, Mandel JS. Colorectal cancer deaths as determined by expert committee and from death certificate: a comparison. The Minnesota Study. *J Clin Epidemiol.* 1999;52(5):447–452.
34. Doria-Rose VP, Marcus PM. Death certificates provide an adequate source of cause of death information when evaluating lung cancer mortality: an example from the Mayo Lung Project. *Lung Cancer.* 2009;63(2):295–300.
35. Freedman DM, Sigurdson AJ, Doody MM, Love-Schnur S, Linet MS. Comparison between cancers identified by state cancer registry, self-report, and death certificate in a prospective cohort study of US radiologic technologists. *Int J Epidemiol.* 2006;35(2):495–497.
36. Barzilai DA, Koroukian SM, Neuhauser D, Cooper KD, Rimm AA, Cooper GS. The sensitivity of medicare data for identifying incident cases of invasive melanoma (United States). *Cancer Causes Control.* 2004;15(2):179–184.
37. Penberthy L, McClish D, Manning C, Retchin S, Smith T. The added value of claims for cancer surveillance: results of varying case definitions. *Med Care.* 2005;43(7):705–712.
38. Cooper GS, Yuan Z, Stange KC, Dennis LK, Amini SB, Rimm AA. The sensitivity of Medicare claims data for case ascertainment of six common cancers. *Med Care.* 1999;37(5):436–444.

Funding

This research was funded in part by the Intramural Research Program of the National Cancer Institute, National Institutes of Health, and by Westat, Inc. (N01-PC-95039 to A.O.G. and M.D.).

Notes

We would like to thank Anthony Polednak, PhD, formerly at the Connecticut Tumor Registry, for his assistance with the registry, and all the respondents and their relatives for participating in the study. The National Cancer Institute was actively involved in the design of the study and collection of data, but the authors had full responsibility for the analysis and interpretation of the data, the decision to submit the article for publication, and the writing of the article.

Affiliations of authors: Division of Cancer Epidemiology and Genetics (PLM, BLG, MHGa, MHGr) and Division of Cancer Control and Population Sciences (AOG, GW), National Cancer Institute, National Institutes of Health, Bethesda, MD; Westat Inc., Rockville, MD (MD); Information Management Services Inc., Silver Spring, MD (TSM); Connecticut Tumor Registry, Hartford, CT (LG); Division of Epidemiology, Services and Prevention Research, National Institute on Drug Abuse, National Institutes of Health, Bethesda, MD (LW).