American Journal of Epidemiology

Vol. 181, No. 3
DOI: 10.1093/aje/kwu258
Advance Access publication:
January 7, 2015

Practice of Epidemiology

Validation of Family Cancer History Data in High-Risk Families: The Influence of Cancer Site, Ethnicity, Kinship Degree, and Multiple Family Reporters

Parisa Tehranifar*, Hui-Chen Wu, Tom Shriver, Ann J. Cloud, and Mary Beth Terry

* Correspondence to Dr. Parisa Tehranifar, Department of Epidemiology, Mailman School of Public Health, Columbia University, 722 West 168th Street, New York, NY 10032 (e-mail: pt140@columbia.edu).

Initially submitted April 28, 2014; accepted for publication September 2, 2014.

Information on family cancer history (FCH) is often collected for first-degree relatives, but more extensive FCH information is critical for greater accuracy in risk assessment. Using self-reported diagnosis of cancer as the gold standard, we examined differences in the sensitivity and specificity of relative-reported FCH by cancer site, race/ethnicity, language preference, and kinship degree (1,524 individuals from 557 families; average number of relatives per family = 2.7). We evaluated the impact of FCH data collected in 2007–2013 from multiple relatives by comparing mean values and proportions for the number of relatives with any cancer, breast cancer, or ovarian cancer as reported by a single relative and by multiple relatives in the same family. The sensitivity of FCH was lower in Hispanics, Spanish-speaking persons, and third-degree relatives (e.g., for all cancers, sensitivities were 80.7%, 87.4%, and 91.0% for third-, second-, and first-degree relatives, respectively). FCH reported by multiple relatives included a higher number of relatives with cancer than the number reported by a single relative (e.g., mean increase of 1.2 relatives with any cancer), with more relatives diagnosed with any cancer, breast cancer, and ovarian cancer in 52%, 36% and 12% of families, respectively. Collection of FCH data from multiple relatives may provide a more comprehensive picture of FCH and may potentially improve risk assessment and preventive care.

cancer; data collection; epidemiologic methods; family medical history; validation studies

Abbreviations: CI, confidence interval; FCH, family cancer history; NY BCFR, New York site of the Breast Cancer Family Registry; Se, sensitivity; Sp, specificity.

Having a family history of cancer is a significant risk factor for many cancers, including breast cancer (1–5). Data on family cancer history (FCH) are incorporated into many public health and clinical guidelines for cancer prevention, early detection, and treatment. FCH information can be used to stratify cancer risks and identify persons who may benefit from more intensive screening programs or from initiation of screening at younger ages. FCH data are also extensively used in basic and population health research to increase knowledge about disease etiology and prevention and have aided in the discovery of cancer susceptibility genes (6, 7). In addition, familial clustering of cancer reflects nongenetic factors that are commonly shared by family members (4, 8–10). For example, relatives who live together or in close proximity for a considerable portion of their lives may have common physical and social environmental exposures, cultural practices, beliefs, and attitudes, as well as similar long-lasting behavioral habits (11–13). Thus, even with increasing advances in the knowledge of genetic factors and genetic testing, family history continues to convey critically significant risk information that goes beyond genetic susceptibility (14–16). Given the importance of FCH to population health research, public health interventions, and clinical care, investigating the validity of FCH is an essential area of research.

The most commonly used approach to collecting FCH information is to ask one family member (hereafter called a "relative") to report disease status for other relatives. The relative reporting the FCH data is more likely to have gathered this information from other relatives than from medical records. As a result, the accuracy of FCH information in most settings is contingent upon the accuracy of self-reported personal history of cancer, as well as sharing of this information

within families. One way to assess whether FCH data reflect these factors is to compare relative-reported FCH with self-reported personal history of cancer. To date, the majority of research on the validity of FCH has only considered FCH in first-degree relatives and has confirmed FCH by comparison with hospital, cancer, and/or death registry data (17–22).

In addition to inadequate accuracy of FCH data, the completeness of FCH data has important implications for both research and clinical practice. Specifically, while a comprehensive FCH should entail at least 3 generations, in most research and clinical settings, FCH information on disease status is collected only for first-degree relatives. In particular, the collection of FCH data for second- or higher-degree relatives is essential for sex-specific cancers for which familial risk may be transmitted through both maternal and paternal lines but can only be observed in first-degree relatives of a specific sex (e.g., breast or ovarian cancer in the paternal line). Because of unavailability of FCH information on more distant relatives, we know substantially less about the feasibility and accuracy of collecting FCH data for second- and third-degree relatives. Collecting FCH information from multiple relatives may be a reasonable way of improving both the accuracy and the completeness of FCH data, as different relatives may be more knowledgeable about the health histories of different relatives within the family; however, currently little is known about the validity of overall FCH compiled through multiple relatives' reports.

To assess the accuracy and completeness of FCH information, we used extensive FCH data collected by 1 or more family members participating in a research registry of families at high risk of breast and/or ovarian cancer. We examined the validity of participants' cancer status as reported by their relatives in comparison with participants' self-reported cancer diagnoses for all cancer sites (except basal-cell skin cancer) and for breast and ovarian cancer. We chose this comparison rather than comparing reports with cancer history ascertained through medical records, because our intent was to capture the accuracy of FCH information as exchanged within families. We further examined whether the validity of relatives' reports of FCH varied by the reporter's race/ethnicity, primary language, and degree of relationship to the relative for whom cancer information was being provided (kinship degree). Finally, we investigated the impact of using multiple relatives' reports on the completeness of FCH information by comparing FCH data based on reports from only 1 relative in the family with FCH data based on reports from multiple relatives from the same family.

METHODS

Study population

We used data from the New York site of the Breast Cancer Family Registry (NY BCFR), a 6-site international research registry established to promote interdisciplinary research on breast cancer etiology and epidemiology (23). The NY BCFR recruited families at high risk of breast and/or ovarian cancer from clinical and community settings within the New York City metropolitan area. The NY BCFR families met at least 1 of the following criteria: 1 female relative diagnosed with breast or ovarian cancer at 45 years of age or younger, 1 female relative diagnosed with both breast and ovarian cancer, 1 male relative diagnosed with breast cancer, 1 relative with a mutation in the breast cancer 1 gene (*BRCA1*) or breast cancer 2 gene (*BRCA2*), or 2 relatives diagnosed with breast and/or ovarian cancer. The NY BCFR study protocol was approved by the Columbia University Medical Center Internal Review Board, and strict quality controls and safeguards were used to protect confidentiality. All participants provided informed consent prior to data collection.

FCH follow-up

In 2007–2013, we began a comprehensive collection of family history information to verify, expand, and update previously collected data on FCH. In contrast to the baseline FCH data collection, in which only 1 relative provided FCH information, all participants at the time of this follow-up were asked to provide FCH data, and each participant's data were entered and stored separately. For assessing the validity of relative-reported FCH as compared with self-reported cancer history, we included families with at least 1 relative reporting his or her own cancer history and at least 1 relative reporting FCH. Therefore, at least 2 family members had to participate in the follow-up to be included in this validity study. A total of 1,524 participants, representing 557 families, met these criteria (average number of relatives per family = 2.7). Over two-thirds (83.8%) of the self-reported cancers were confirmed through pathology reports. In

Table 1. Characteristics of Participants With Follow-up Data on Family Cancer History (n = 1,524 Individuals), New York Breast Cancer Family Registry, 2007–2013

Characteristic	No.	%
Age, years[a]	56.9 (14.5)[b]	
Sex		
Female	1,221	80.1
Male	303	19.9
Race/ethnicity		
White	1,143	75.0
Hispanic	282	18.5
Other	99	6.5
Language used in data collection		
English	1,448	95.0
Spanish	76	5.0
Personal history of breast cancer		
No	1,066	70.0
Yes	458	30.0
Personal history of ovarian cancer		
No	1,486	97.5
Yes	38	2.5

[a] Age at the time of completion of the follow-up questionnaire.
[b] Value presented as mean (standard deviation).

Table 2. Sensitivity and Specificity of Relatives' Reports of Any Cancer, Female Breast Cancer, and Ovarian Cancer by Cancer Site and Relatives' Race/Ethnicity and Primary Language, New York Breast Cancer Family Registry, 2007–2013

Cancer Site and Report Variable	No. of True-Positive Reports	No. of False-Positive Reports	No. of False-Negative Reports	No. of True-Negative Reports	Sensitivity, %	95% CI	*P* Value	Specificity, %	95% CI	*P* Value
All Cancer Sites										
All reports	1,372	288	175	2,261	88.7	87.0, 90.2		88.7	87.4, 89.9	
Kinship degree										
First-degree relatives	906	163	90	1,464	91.0	89.0, 92.7		90.0	88.4, 91.4	
Second-degree relatives	215	51	31	438	87.4	82.6, 91.3		89.6	86.5, 92.1	
Third-degree relatives	205	66	49	307	80.7	75.3, 85.4	<0.0001[a]	82.3	78.0, 86.0	<0.0001[a], 0.006[b]
Race/ethnicity										
White	1,060	204	154	1,604	87.3	85.3, 89.1		88.7	87.2, 90.1	
Hispanic	243	72	16	549	93.8	90.2, 96.4	0.009[c]	88.4	85.6, 90.8	
Other	69	12	5	108	93.2	84.9, 97.8		90.0	83.2, 94.7	
Language of data collection										
English	1,291	253	167	2,126	88.5	86.8, 90.1		89.4	88.1, 90.6	
Spanish	81	35	8	135	91.0	83.1, 96.0		79.4	72.5, 85.2	<0.0001
Female Breast Cancer										
All reports	1,013	154	68	2,020	93.7	92.1, 95.1		92.9	91.8, 94.0	
Kinship degree										
First-degree relatives	694	75	45	1,301	93.9	91.9, 95.5		94.5	93.2, 95.7	
Second-degree relatives	155	16	7	388	95.7	91.3, 98.2		96.0	93.7, 97.7	
Third-degree relatives	141	61	15	304	90.4	84.6, 94.5		83.3	79.1, 87.0	<0.0010[a], <0.0001[b]
Race/ethnicity										
White	767	135	39	1,366	95.2	93.4, 96.5		91.0	89.4, 92.4	
Hispanic	190	13	26	561	88.0	82.9, 92.0	0.0004[c]	97.7	96.2, 98.8	<0.0001[c]
Other	56	6	3	93	94.9	85.9, 98.9		93.9	87.3, 97.7	
Language of data collection										
English	954	150	55	1,857	94.5	93.0, 95.9		92.5	91.3, 93.6	
Spanish	59	4	13	163	81.9	71.1, 90.0	<0.0001	97.6	94.0, 99.3	0.01
Ovarian Cancer										
All reports	88	18	22	3,127	80.0	71.3, 87.0		99.4	99.1, 99.7	
Kinship degree										
First-degree relatives	57	10	9	2,039	86.4	75.7, 93.6		99.5	99.1, 99.8	
Second-degree relatives	13	4	8	541	61.9	38.4, 81.9	0.04[a]	99.3	98.1, 99.8	

Table continues

Table 2. Continued

Cancer Site and Report Variable	No. of True-Positive Reports	No. of False-Positive Reports	No. of False-Negative Reports	No. of True-Negative Reports	Sensitivity, %	95% CI	*P* Value	Specificity, %	95% CI	*P* Value
Third-degree relatives	15	4	5	497	75.0	50.9, 91.3		99.2	98.0, 99.8	
Race/ethnicity										
White	73	13	17	2,204	81.1	71.5, 88.6		99.4	99.0, 99.7	
Hispanic	9	3	4	774	69.2	38.6, 90.9		99.6	98.9, 99.9	
Other	6	2	1	149	85.7	42.1, 99.6		98.7	95.3, 99.8	
Language of data collection										
English	83	17	20	2,896	80.6	71.6, 87.7		99.4	99.1, 99.7	
Spanish	5	1	2	231	71.4	29.0, 96.3		99.6	97.6, 100	

Abbreviation: CI, confidence interval.
[a] Bonferroni-adjusted *P* value for comparison with first-degree relatives.
[b] Bonferroni-adjusted *P* value for comparison with second-degree relatives.
[c] Bonferroni-adjusted *P* value for comparison with whites.

addition to providing information on personal history of cancer, participants reported the following information on their living and deceased relatives: first name, date of birth, date of death, presence or absence of a cancer diagnosis, and cancer site if a diagnosis was reported. Degrees of kinship were defined as follows: first-degree relatives were parents, siblings, and offspring; second-degree relatives were grandparents, grandchildren, aunts, uncles, nieces, and nephews; and third-degree relatives were cousins, great-grandparents, and great-grandchildren. For comparison of FCH information provided by a single relative with that provided by multiple relatives, we used FCH data from 546 families with at least 2 relatives per family providing FCH data.

Statistical methods

We calculated sensitivity and specificity and their corresponding 95% confidence intervals to compare cancer status data reported by the participants themselves (self-reports) with cancer status data reported for them by their relatives (relative reports). In our analysis, sensitivity was the proportion of self-reported cancer diagnoses that was correctly classified by relative reports, and specificity was the proportion of negative self-reported cancer diagnoses that was correctly classified by relative reports. These calculations were performed for reports of any cancer, breast cancer, and ovarian cancer, and the analyses were stratified according to relatives' characteristics, including race/ethnicity (white, Hispanic, other), primary language used for data collection (English, Spanish), and degree of kinship with the person for whom cancer status was being reported (first-, second-, or third-degree relative). We used χ^2 tests for proportions, with Bonferroni adjustment for multiple comparisons, to evaluate the statistical significance of observed differences across comparison groups. To evaluate the impact of having FCH data provided by multiple relatives, we compared the number of family members diagnosed with any cancer, breast cancer, or ovarian cancer as reported by all relatives in the family with the number of family members with the same type(s) of cancer as reported by 1 relative in the same family. The comparison was with a single reporter from the family who was unaffected by breast cancer. For these analyses, the family boundaries remained the same when comparing index relatives' and multiple relatives' FCH reports. All reported *P* values are 2-sided, and all statistical tests were performed using SAS 9.2 (SAS Institute, Inc., Cary, North Carolina).

RESULTS

Participants included in this analysis were representative of the NY BCFR population, with the majority being female (80%), non-Hispanic white (75%), and English speakers (95%). Approximately 30% had a history of breast cancer, and fewer than 3% had a history of ovarian cancer (Table 1).

We evaluated the accuracy of relative-reported FCH in comparison with self-reported cancer by race/ethnicity, kinship degree, and primary language for all cancer sites and for breast and ovarian cancers (Table 2). Sensitivity was highest for breast cancer status (sensitivity (Se) = 93.7%, 95% confidence interval (CI): 92.1, 95.1) and lowest for ovarian cancer

status (Se = 80.0%, 95% CI: 71.3, 87.0). The sensitivity for all cancer sites showed intermediate values between those for breast cancer and those for ovarian cancer (Se = 88.7, 95% CI: 87.0, 90.2), and specificity for all cancer sites was lower than specificities for both breast and ovarian cancer (specificity (Sp) = 88.7, 95% CI: 87.4, 89.9).

The accuracy of FCH data showed some variation by kinship degree for all cancer sites and for ovarian cancer. For example, for all cancer sites, sensitivity was significantly lower for third-degree relative reports (Se = 80.7%, 95% CI: 75.3, 85.4) than for first-degree relative reports (Se = 91.0, 95% CI: 89.0, 92.7), and specificity was significantly lower for third-degree relative reports (Sp = 82.3%, 95% CI: 78.0, 86.0) than for first-degree relative (Sp = 90.0%, 95% CI: 88.4, 91.4) and second-degree relative (Sp = 89.6%, 95% CI: 86.5, 92.1) reports (Table 2). The accuracy of FCH data for all cancer sites and for breast cancer also differed significantly by race/ethnicity and language, with the lowest sensitivity for breast cancer being observed among participants who were Hispanic (Se = 88.0%, 95% CI: 82.9, 92.0) and Spanish-speaking (Se = 81.9%, 95% CI: 71.1, 90.0). Hispanic and Spanish-speaking participants also had the highest specificity for breast cancer (Sp = 97.7% (95% CI: 96.2, 98.8) and Sp = 97.6 (95% CI: 94.0, 99.3), respectively). In contrast, Spanish-speaking participants had the lowest specificity for all cancer sites (Sp = 79.4%, 72.5, 85.2). With the exception of lower sensitivities in third-degree relatives and second-degree relatives relative to first-degree relatives, the accuracy of FCH data on ovarian cancer did not show statistically significant variations by other factors. We also examined the associations presented in Table 2 using only 1 reporter per family member; results were very similar, and the overall trends were the same (data not shown).

We further examined sensitivity and specificity among first-degree relatives according to the type of relationship relatives had to the person for whom cancer status data were being provided (i.e., parents, siblings, and offspring) (Table 3). The only statistically significant differences observed were for the specificity of all cancer sites, with reports from parents having lower specificity (Sp = 81.4%, 95% CI: 76.2, 85.9) than reports from siblings (Sp = 90.5%, 95% CI: 88.5, 92.3) and offspring (Sp = 94.3%, 95% CI: 91.6, 96.4).

To assess the completeness of FCH information, we compared the reports of cancer at any site, breast cancer, and ovarian cancer made by the index relatives with reports of the same information made by all participating relatives in the same family (multiple relatives), keeping the same family boundaries. All comparisons showed statistically significantly higher mean numbers of relatives with a cancer history when FCH was reported by multiple relatives in the family than when FCH was reported by the index relative, with mean increases of 1.2, 0.6, and 0.1 in the numbers of relatives with cancer at all sites, breast cancer, and ovarian cancer, respectively (Table 4). As compared with FCH data reported by the index relative, multiple relatives collectively reported more relatives diagnosed with any cancer, breast cancer, and ovarian cancer in 52%, 36%, and 12% of 546 families with 2 or more relatives reporting FCH, respectively. The proportion of increase in the number of cancer cases within the family increased as the number of multiple relatives reporting FCH increased. For example, 16% of families with 2 relative reporters reported more breast cancer cases, whereas 65% of families with 4 or more relative reporters reported more breast cancer cases, both in comparison with reports made by the index relatives in the same families (Figure 1B).

DISCUSSION

FCH reflects the influences of shared environments, genes, and behaviors among relatives, and if the information is collected accurately, it can provide critical information for etiological and prevention research and for early detection, risk reduction, and treatment interventions. In recent years, growing attention has been focused on improving the collection and validity of family health history data, including the establishment of the Family Health History Initiative (http://www.hhs.gov/familyhistory) by the Office of the Surgeon General (US Public Health Service) to raise public awareness of the importance of family health information and to facilitate communication and collection of this information by families (24–26). To understand the validity of FCH information communicated within families, we systematically collected FCH data from multiple relatives and compared relative reports of cancer status with self-reports of the same information among families at high risk for breast or ovarian cancer. We found high levels of sensitivity and specificity for family history of breast cancer and to a lesser extent for all cancer sites combined. The sensitivity of family history of ovarian cancer was more moderate, while specificity for this disease was extremely high. We observed more significant variations in specificity than in sensitivity measures; most notably, significantly lower specificity was found for reports by third-degree relatives as compared with first- and second-degree relatives for all cancer sites and breast cancer. Sensitivity was also lower for second- and third-degree relatives for all cancer sites and for ovarian cancers, but there was little variation in the sensitivity of breast cancer status across different degrees of kinship. Furthermore, the accuracy of FCH information on breast cancer and all cancer sites was lower among Hispanic and Spanish-speaking participants, although the sensitivity remained high at >80%. Overall, these findings suggest that even in high-risk and highly motivated families participating in family-based studies, FCH is subject to systematic variations by kinship degree and sociodemographic variables, as reported in prior research with average-risk populations (17, 18, 21, 27).

The sensitivities of FCH data in our study were close to the highest range of sensitivities reported in other studies, with most researchers reporting sensitivity measures in the range of 33%–95% (reviewed by Qureshi et al. (22)). In a recent population-based study of over 1,000 participants that considered kinship degree, Mai et al. (18) reported sensitivities of 64.9% and 59.0% for breast cancer family history reported by first-degree relatives and second-degree relatives, respectively, which are also considerably lower than the values we observed in this study across kinship levels. Our study population included families with at least 1 member diagnosed with breast or ovarian cancer or identified as a carrier of a mutation in *BRCA1* or *BRCA2*. The majority of the families also had multiple relatives participating in the registry. Given

Table 3. Sensitivity and Specificity of First-Degree Relatives' Reports of Any Cancer, Female Breast Cancer, and Ovarian Cancer by Kinship Type, New York Breast Cancer Family Registry, 2007–2013

Cancer Site and Type of First-Degree Relative	No. of True-Positive Reports	No. of False-Positive Reports	No. of False-Negative Reports	No. of True-Negative Reports	Sensitivity, %	95% CI	Specificity, %	95% CI
Any Cancer								
Parents	279	49	21	215	93.0	89.5, 95.6	81.4[a,b]	76.2, 85.9
Siblings	498	91	58	866	89.6	86.7, 92.0	90.5	88.5, 92.3
Offspring	129	23	11	383	92.1	86.4, 96.0	94.3	91.6, 96.4
Female Breast Cancer								
Parents	190	11	12	203	94.1	89.9, 96.9	94.9	91.0, 97.4
Siblings	396	47	25	777	94.1	91.4, 96.1	94.3	92.5, 95.8
Offspring	108	17	8	321	93.1	86.9, 97.0	95.0	92.1, 97.0
Ovarian Cancer								
Parents	17	2	1	396	94.4	72.7, 99.9	99.5	98.2, 99.9
Siblings	33	7	7	1,198	82.5	67.2, 92.7	99.4	98.8, 99.8
Offspring	7	1	1	445	87.5	47.4, 99.7	99.8	98.7, 99.9

Abbreviation: CI, confidence interval.
[a] $P < 0.05$ (Bonferroni-adjusted P value for comparison with siblings).
[b] $P < 0.05$ (Bonferroni-adjusted P value for comparison with offspring).

these study design characteristics, it is reasonable to assume that there is a fairly high degree of sharing of FCH information within families participating in the BCFR, which may be reflected in the higher sensitivity of FCH data in our study as compared with other studies. Furthermore, with the exception of FCH data for ovarian cancer, the ranges of sensitivity and specificity values were similar in our study, whereas in other studies, the specificity of FCH data tended to be considerably higher than the sensitivity. These results may suggest that among high-risk families, overreporting of cancer status may be as much of a concern as underreporting of FCH.

We were also interested in learning whether collection of FCH data from additional family members would yield a more complete picture of FCH. To this end, we compared FCHs reported by a single relative with FCHs in the same family that were reported by multiple relatives. This comparison was made to mimic what happens in a typical epidemiologic cohort when unaffected women may be asked about their family history. Using the same family size (boundaries) for these comparisons of single reporters with multiple reporters, we found an increased number of relatives with cancer in FCHs based on multiple relatives versus a single relative. Given the observed high accuracy of relative reports of cancer, the larger number of relatives with cancer obtained using multiple relatives' reports suggests that inclusion of more relatives in the collection of FCH data identified cancer cases in the family that may have been missed when relying on a single relative informant in the family. For example, in one family, FCH data obtained from multiple relatives identified 2 additional family members with breast cancer who had not been included in the index relative's report of FCH. Based on the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) model (http://ccge.medschl.cam.ac.uk/boadicea/), the additional breast cancer cases in the family increased the estimated remaining lifetime risk of a 53-year-old female family member to greater than 20%, the clinical threshold for breast cancer screening and prevention guidelines (28, 29). The additional FCH reported by multiple relatives can also lead to lower risk profiles based on new information on key variables, such as more advanced age

Table 4. Number of Relatives Diagnosed With Cancer as Reported by 1 (Index) Relative and as Reported by Multiple Relatives Within the Same Family, by Cancer Site ($n = 546$ Families), New York Breast Cancer Family Registry, 2007–2013

Cancer Site	Index Relative Reporter		Multiple Relative Reporters		Mean Difference	95% Confidence Interval
	Mean (SD)	Range	Mean (SD)	Range		
Any cancer	2.5 (1.7)	0–10	3.8 (2.5)	0–17	1.2	1.1, 1.4
Female breast cancer	1.3 (1.1)	0–8	1.9 (1.4)	0–8	0.6	0.5, 0.6
Ovarian cancer	0.2 (0.6)	0–3	0.3 (0.6)	0–5	0.1	0.1, 0.2

Abbreviation: SD, standard deviation.

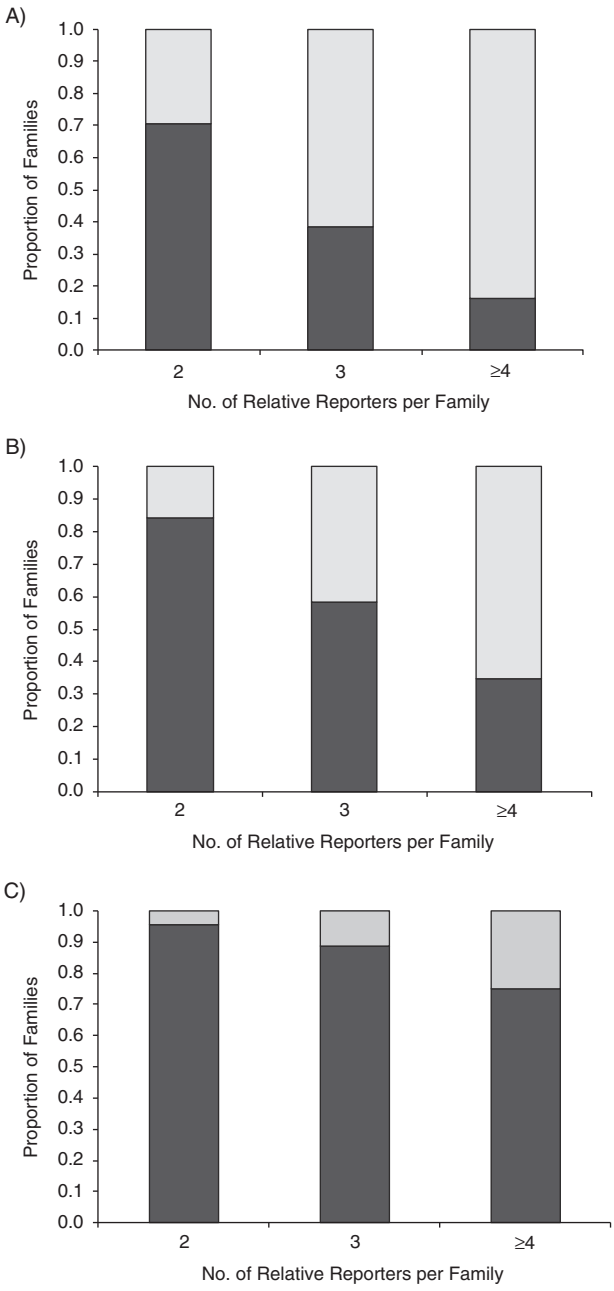

Figure 1. Differences in the numbers of family members diagnosed with cancer at any site (A), breast cancer (B), and ovarian cancer (C), as reported by a single relative (index) reporter and by multiple relative reporters, according to number of relatives reporting family history data ($n = 546$ families), New York Breast Cancer Family Registry, 2007–2013. Light shading shows the number of families with more family members diagnosed with cancer as reported by multiple relatives versus reported by a single relative in the family; dark shading shows the number of families with no differences in the number of family members diagnosed with cancer as reported by multiple relatives versus reported by a single relative in the family.

at cancer diagnosis by relatives with a closer relationship to affected family members. In our study, 2 or more relatives contributed FCH data in over 90% of the families, suggesting that collection of data from multiple family members may be practical and feasible in certain settings, such as family-based studies.

We used self-reported cancer diagnosis as the "gold standard" to measure the sensitivity and specificity of relative-reported FCH. We recognize that the true gold standard for cancer should be pathology reports; however, given that family members' knowledge of their FCH is most often obtained from a relative's self-report of personal cancer history, this type of validity information has important practical implications in many settings. The literature on the validity of self-reported cancer as compared with cancer registry data generally shows moderate-to-high accuracy, with higher accuracy for certain cancers (e.g., breast cancer) and among persons participating in cancer-related projects (27, 30–34). Participants in our registry represent a relatively wide spectrum of cancer risk, ranging from being a *BRCA1/2* mutation carrier to having only 1 relative in the family with breast or ovarian cancer; however, the great majority of study participants are likely to be at greater risk for breast or ovarian cancer than the general population. While this study design feature limits the external generalizability of our results to other populations, it provides an opportunity to demonstrate the higher range of accuracy for relative-reported FCH. This has important implications for accurately capturing family history of ovarian cancer, with only 86% of all cases being correctly reported even by first-degree relatives. Because we restricted our analysis to participants with self-reported cancer status at follow-up, our study population is likely to have included a higher proportion of cancer survivors, particularly for fatal cancers such as ovarian cancer; however, it is not clear whether this limitation would increase or reduce the validity of FCH information.

The main strengths of our study included the use of a large sample size, a high participation rate, and unique FCH data, which were systematically collected from multiple participants within the same family. Together, these strengths allowed for a comprehensive examination of the accuracy and completeness of relative-reported FCH.

In summary, families at high risk of breast or ovarian cancer have relatively accurate knowledge of their FCH, but this information is less accurately reported by Hispanic and Spanish-speaking persons and by third-degree relatives. Efforts to improve communication about FCH within families have the potential to increase accurate reporting of this crucial information, with important implications for both research and clinical practice. Our results also suggest that the addition of multiple relatives in the collection of FCH data, if feasible, may improve the completeness and accuracy of FCH by capturing missed cases of cancer and may provide additional or more accurate details on cancer history, such as ages of diagnosis. A more complete and accurate portrait of FCH is critical to accurate risk assessment for guidance on preventive care decisions, including chemoprevention and risk-reducing surgeries.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York (Parisa Tehranifar, Hui-Chen Wu, Tom Shriver, Ann

J. Cloud, Mary Beth Terry); Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, New York, New York (Parisa Tehranifar, Mary Beth Terry); and Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York (Hui-Chen Wu).

This work was supported by an award from the Breast Cancer Research Foundation and by National Institutes of Health grants U01 CA69398, P30 CA13696, and P30 ES009089. This work was also supported by the National Cancer Institute under grant CA-06-503 and through cooperative agreements with members of the Breast Cancer Family Registry (BCFR) and principal investigators. The first author (P.T.) was supported by National Cancer Institute grant K07 CA151777.

The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the BCFR, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government or the BCFR.

Conflict of interest: none declared.

REFERENCES

1. Whittemore AS, Wu AH, Kolonel LN, et al. Family history and prostate cancer risk in black, white, and Asian men in the United States and Canada. *Am J Epidemiol*. 1995;141(8):732–740.
2. Hunn J, Rodriguez GC. Ovarian cancer: etiology, risk factors, and epidemiology. *Clin Obstet Gynecol*. 2012;55(1):3–23.
3. McWilliams RR, Rabe KG, Olswold C, et al. Risk of malignancy in first-degree relatives of patients with pancreatic carcinoma. *Cancer*. 2005;104(2):388–394.
4. Pharoah PD, Day NE, Duffy S, et al. Family history and the risk of breast cancer: a systematic review and meta-analysis. *Int J Cancer*. 1997;71(5):800–809.
5. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. *Am J Gastroenterol*. 2001;96(10):2992–3003.
6. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*. 1994;266(5182):66–71.
7. Wooster R, Neuhausen SL, Mangion J, et al. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science*. 1994;265(5181):2088–2090.
8. Carpenter CL, Ross RK, Paganini-Hill A, et al. Effect of family history, obesity and exercise on breast cancer risk among postmenopausal women. *Int J Cancer*. 2003;106(1):96–102.
9. Chen PL, Sellers TA, Rich SS, et al. Examination of the effect of nongenetic risk factors on the familial risk of breast cancer among relatives of postmenopausal breast cancer patients. *Cancer Epidemiol Biomarkers Prev*. 1994;3(7):549–555.
10. Ortega-Alonso A, Sipilä S, Kujala UM, et al. Genetic influences on adult body mass index followed over 29 years and their effects on late-life mobility: a study of twin sisters. *J Epidemiol Community Health*. 2009;63(8):651–658.
11. Mitchell BD, Rainwater DL, Hsueh WC, et al. Familial aggregation of nutrient intake and physical activity: results from the San Antonio Family Heart Study. *Ann Epidemiol*. 2003;13(2):128–135.
12. Oliveria SA, Ellison RC, Moore LL, et al. Parent-child relationships in nutrient intake: the Framingham Children's Study. *Am J Clin Nutr*. 1992;56(3):593–598.
13. Jacobi D, Caille A, Borys JM, et al. Parent-offspring correlations in pedometer-assessed physical activity. *PLoS One*. 2011;6(12):e29195.
14. Yoon PW, Scheuner MT, Peterson-Oehlke KL, et al. Can family history be used as a tool for public health and preventive medicine? *Genet Med*. 2002;4(4):304–310.
15. Guttmacher AE, Collins FS, Carmona RH. The family history—more important than ever. *N Engl J Med*. 2004;351(22):2333–2336.
16. Ramsey SD, Yoon P, Moonesinghe R, et al. Population-based study of the prevalence of family history of cancer: implications for cancer screening and prevention. *Genet Med*. 2006;8(9):571–575.
17. Ziogas A, Anton-Culver H. Validation of family history data in cancer family registries. *Am J Prev Med*. 2003;24(2):190–198.
18. Mai PL, Garceau AO, Graubard BI, et al. Confirmation of family cancer history reported in a population-based survey. *J Natl Cancer Inst*. 2011;103(10):788–797.
19. Ozanne EM, O'Connell A, Bouzan C, et al. Bias in the reporting of family history: implications for clinical care. *J Genet Couns*. 2012;21(4):547–556.
20. Qureshi N, Wilson B, Santaguida P, et al. Collection and use of cancer family history in primary care. *Evid Rep Technol Assess (Full Rep)*. 2007;(159):1–84.
21. 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.
22. Qureshi N, Wilson B, Santaguida P, et al. *NIH State-of-the-Science Conference: Family History and Improving Health*. Rockville, MD: Agency for Healthcare Research and Quality; 2009. (AHRQ publication no. 09-E016).
23. John EM, Hopper JL, Beck JC, et al. The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. *Breast Cancer Res*. 2004;6(4):R375–R389.
24. 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.
25. Valdez R, Yoon PW, Qureshi N, et al. Family history in public health practice: a genomic tool for disease prevention and health promotion. *Annu Rev Public Health*. 2010;31:69–87.
26. Rich EC, Burke W, Heaton CJ, et al. Reconsidering the family history in primary care. *J Gen Intern Med*. 2004;19(3):273–280.
27. Stavrou E, Vajdic CM, Loxton D, et al. The validity of self-reported cancer diagnoses and factors associated with accurate reporting in a cohort of older Australian women. *Cancer Epidemiol*. 2011;35(6):e75–e80.
28. National Comprehensive Cancer Network, Inc. *Breast Cancer Screening and Diagnosis*. (Practice Guidelines in Oncology, version 1.2010). Fort Washington, PA: National Comprehensive Cancer Network, Inc.; 2009.
29. Saslow D, Boetes C, Burke W, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin*. 2007;57(2):75–89.
30. Parikh-Patel A, Allen M, Wright WE, et al. Validation of self-reported cancers in the California Teachers Study. *Am J Epidemiol*. 2003;157(6):539–545.
31. Desai MM, Bruce ML, Desai RA, et al. Validity of self-reported cancer history: a comparison of health interview data and cancer registry records. *Am J Epidemiol*. 2001;153(3):299–306.

32. Schrijvers CT, Stronks K, van de Mheen DH, et al. Validation of cancer prevalence data from a postal survey by comparison with cancer registry records. *Am J Epidemiol*. 1994;139(4):408–414.
33. Bergmann MM, Calle EE, Mervis CA, et al. Validity of self-reported cancers in a prospective cohort study in comparison with data from state cancer registries. *Am J Epidemiol*. 1998;147(6):556–562.
34. Manjer J, Merlo J, Berglund G. Validity of self-reported information on cancer: determinants of under- and over-reporting. *Eur J Epidemiol*. 2004;19(3):239–247.