

Self-report versus Medical Records for Assessing Cancer-Preventive Services Delivery

Jeanne M. Ferrante,^{1,2,3} Pamela Ohman-Strickland,^{1,2,4} Karissa A. Hahn,^{1,2} Shawna V. Hudson,^{1,2} Eric K. Shaw,¹ Jesse C. Crosson,^{1,3} and Benjamin F. Crabtree^{1,2,5}

¹Department of Family Medicine and ²The Cancer Institute of New Jersey, University of Medicine and Dentistry, New Jersey (UMDNJ)-Robert Wood Johnson Medical School, New Brunswick, New Jersey; ³Department of Family Medicine, UMDNJ-New Jersey Medical School, Newark, New Jersey; and Departments of ⁴Biostatistics and ⁵Epidemiology, UMDNJ-School of Public Health, Piscataway, New Jersey

Abstract

Accurate measurement of cancer-preventive behaviors is important for quality improvement, research studies, and public health surveillance. Findings differ, however, depending on whether patient self-report or medical records are used as the data source. We evaluated concordance between patient self-report and medical records on risk factors, cancer screening, and behavioral counseling among primary care patients. Data from patient surveys and medical records were compared from 742 patients in 25 New Jersey primary care practices participating at baseline in SCOPE (supporting colorectal cancer outcomes through participatory enhancements), an intervention trial to improve colorectal cancer screening in primary care offices. Sensitivity, specificity, and rates of agreement describe concordance between self-report and medical records for risk factors (personal or family history of cancer, smoking), cancer screening

(breast, cervical, colorectal, prostate), and counseling (cancer screening recommendations, diet or weight loss, exercise, smoking cessation). Rates of agreement ranged from 41% (smoking cessation counseling) to 96% (personal history of cancer). Cancer screening agreement ranged from 61% (Pap and prostate-specific antigen) to 83% (colorectal endoscopy) with self-report rates greater than medical record rates. Counseling was also reported more frequently by self-report (83% by patient self-report versus 34% by medical record for smoking cessation counseling). Deciding which data source to use will depend on the outcome of interest, whether the data is used for clinical decision making, performance tracking, or population surveillance; the availability of resources; and whether a false positive or a false negative is of more concern. (Cancer Epidemiol Biomarkers Prev 2008;17(11):2987–94)

Introduction

Accurate measurement of cancer-preventive behaviors is important for assessing quality of medical care, conducting descriptive and interventional research studies, monitoring progress toward national goals, and determining public policies. Findings differ, however, depending on whether administrative data, patient self-report, or medical records are used as the data source. Data from insurance claims and medical records have been found to have high concordance for certain procedures such as colonoscopy (1); however, relatively lower levels of agreement between medical record review and patient self-report have been found in the documentation of diabetes care, cardiovascular disease care, cancer screening, and counseling or referral to additional services (2–4). Further, the extent of this disagreement has been found to vary by type of practice organization (5).

The best source for information to assess the quality of cancer-preventive care in primary care settings remains uncertain. Insurance claims data may be most accurate for procedures that generate a bill (1, 3); however, using claims data is impractical in community primary care settings that have patients from multiple insurance plans. Therefore, medical records are frequently used as the data source for assessing quality of medical care and conducting health services research in community settings. Medical records are relatively accessible, presumed to have accurate clinical information, and provide clinical richness (6). However, medical record review is costly and labor- and time-intensive, especially when manually reviewing records one by one, and it depends on the completeness and accuracy of documentation and abstraction (5, 6). The quality of medical record documentation may be affected by misfiling of reports or charts, tests done elsewhere and not reported back to the practice, less documentation by busy physicians, delayed recording by physicians leading to errors in recall, and illegibility. Furthermore, because medical records are constructed for legal or billing reasons as well as for documenting clinical information, unbillable procedures such as those associated with counseling and some screenings are less likely to be recorded (7, 8). In addition to problems with recording data in the charts,

Received 2/22/08; revised 7/28/08; accepted 8/6/08.

Grant support: National Cancer Institute grant 5 R01 NCI CA11287.

Requests for reprints: Jeanne M. Ferrante, Department of Family Medicine, University of Medicine and Dentistry, New Jersey-Robert Wood Johnson Medical School, 1 World's Fair Drive, Suite 1515, Somerset, NJ 08873. Phone: 732-743-3386; Fax: 732-743-3395. E-mail: ferranjm@umdnj.edu

Copyright © 2008 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-08-0177

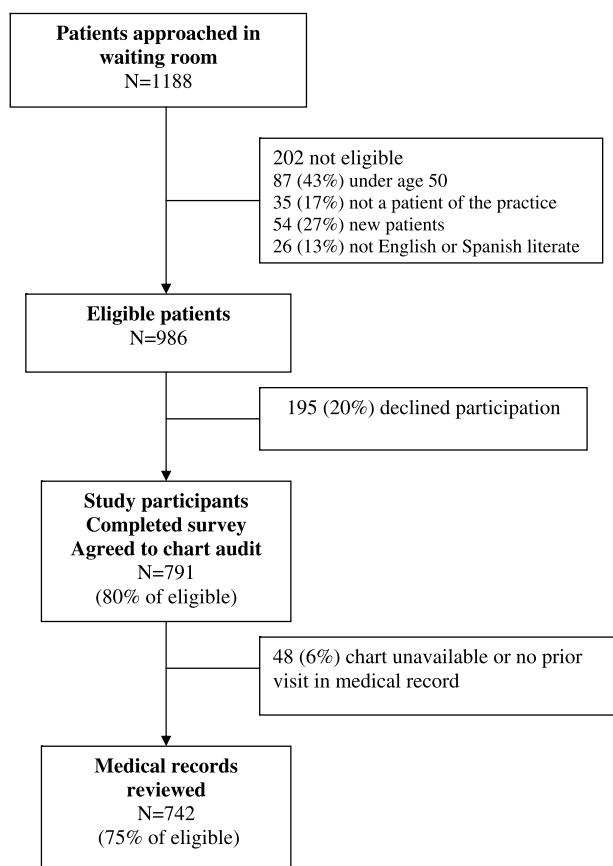


Figure 1. Subject flow diagram.

limitations of the reviewer (errors from fatigue, inattentiveness, poor training, and systematic reviewer bias) may affect the accuracy of medical record abstraction (6).

Using patient self-report data may be relatively cheaper, faster, and easier than manually extracting data from medical record review, but there may also be limitations in its quality, due to bias in patient recall, patient's lack of knowledge of screening tests, poorly designed survey instruments, and untruthful responses due to social desirability bias (3). In studies focused on documenting the receipt of cancer-related services, patients have been found to inaccurately report family histories of specific cancers (9, 10), underestimate the time since last breast or cervical cancer screening (11-14), mistakenly report cervical cancer screening when receiving pelvic exams for other reasons (15), overreport receipt of colorectal (1, 16-18) or prostate cancer screening (17, 19, 20), and mistakenly report the type of colorectal cancer screening (21). However, patient self-report has been found to accurately determine changes in screening status over time (22), a past history of cancer (23), or smoking status (24). Nonetheless, relying on patient self-report typically yields significantly higher estimates of cancer screening than relying on medical records (3, 18, 25, 26).

Several studies have compared cancer-related data from medical record review and patient self-report, but few addressed which source is preferred. Many studies

assumed medical records as the gold standard (11-13, 16-20, 26-30), included only patients enrolled in health maintenance organizations (11, 12, 16, 17, 28, 31), patients from one clinical setting (20), an integrated health care system (32, 33), non-primary care clinics (13), or inner city settings (15, 27, 34), and only addressed one or two specific cancer screenings (1, 11-13, 16, 17, 19, 20, 27, 29-33, 35). The few studies examining the validity of self-reported personal or family history of cancer compared self-report to cancer registries (9, 10, 23, 36-39), and most were from non-U.S. settings (9, 10, 36, 37). Studies addressing which data source is preferred had other limitations. Stange et al. compared medical records and patient self-report with direct observation of several clinical services in primary care practices; however, the patient interviews were collected immediately after visits, which may lead to fewer recall errors than in more typical patient surveys (8). Tisnado et al. examined levels of concordance and variations for risk factors and behavioral counseling, but they did not examine cancer screening tests (4). Conversely, Montañó and Phillips evaluated correlation between the two data sources on cancer screening tests, but they did not address risk factors or behavioral counseling (40). We could find no studies that compared recommendations for cancer screening between patient self-report and medical records.

The purpose of this study was to examine concordance between patient self-report and medical record review in determining cancer-related risk factors, screening tests, and behavioral counseling among patients in primary care practices. Recognizing that neither method is perfect, we also describe the relative performance of each data source by calculating the sensitivity and specificity of report of each item using both medical record review and patient self-report as a reference.

Materials and Methods

We used cross-sectional data collected at baseline, from January 2006 through May 2007, as part of a quality improvement intervention study, Supporting Colorectal Cancer Outcomes through Participatory Enhancements (SCOPE). The SCOPE study used a multimethod assessment process (41) to inform a facilitated team-building intervention (42) aimed at improving colorectal cancer screening among 25 practices in the New Jersey Family Medicine Research Network. The University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School Institutional Review Board approved this study. Written informed consent to participate in the study was received from the medical directors and/or lead physicians of each practice as well as from patients and staff members who participated in the study.

Data Collection. Data were collected via patient surveys and medical record review. A convenience sample of 30 consecutive patients, age 50 years or older, were recruited in the waiting rooms of each practice. New patients and those who could not read or write in English or Spanish were excluded. After informed consent, all patients completed a self-administered written survey that included demographics, risk factors,

dates of cancer screenings, health care seeking behaviors, self-rated health, and satisfaction with care. For patients who refused to participate, recruiters recorded their gender and approximate age based on appearance. Participants and those who refused were similar in gender, but they differed in age, with older patients more likely to participate than younger patients. For example, 68% of eligible patients ages 50 to 59 years agreed to participate, whereas 90% of those ages 60 to 69 years and 85% of those age 70 years and older agreed to participate ($P < 0.001$). Each patient also consented to have their medical record reviewed. Eighty percent ($n = 791$) of eligible patients approached in the waiting room completed the patient survey and agreed to have their medical record reviewed. We excluded 48 patients whose chart was unavailable or who did not have at least one prior visit in the medical record. Therefore, complete survey and medical record data were available for 742 patients (see Fig. 1).

Trained nurse chart auditors used a standardized chart abstraction tool to record patient age, weight, cancer risk factors, medical conditions, dates of cancer screenings and recommendations, and health behavior counseling. In addition to the medical record review, practice managers and lead physicians completed a 46-item practice information survey that obtained information regarding the practice.

Outcome Measures. Our main outcomes of interest were documented or reported evidence of cancer-related risk factors, cancer screening, and behavioral counseling.

1. Cancer-related risk factors: All patients were assessed for personal history of cancer, family history of cancer, and smoking status. Personal history and family history of cancer included breast, cervical, colorectal, or prostate cancer.
2. Cancer screening: Patients were asked if they received breast (women only), cervical (women only), colorectal (all patients), or prostate (men only) cancer screening and for the dates of their most recent mammogram, Pap smear, fecal occult blood test (FOBT), sigmoidoscopy, colonoscopy, or prostate-specific antigen (PSA) test. The self-report questions on cancer screening were taken from the 2005 National Health Interview Survey questionnaire (43). Patients were asked if they ever had the cancer screening test, followed by a description of the test as in National Health Interview Survey. Those who responded "yes" were asked if they had the test within the guideline specified time period (i.e., 1 year for mammography), as well as the date of their last test (month, year). Medical record documentation of cancer screening was obtained by searching all aspects of the medical record including progress reports, preventive flow sheets, laboratory tests, X-rays, and consultant reports. Up-to-date in cancer screening was determined based on recommendations from the American Cancer Society: (a) breast cancer (mammogram within 1 year); (b) cervical cancer (Pap smear within 3 years); (c) colorectal cancer (FOBT within 1 year, sigmoidoscopy or barium enema within 5 years, or colonoscopy within 10 years); and (d) prostate cancer (PSA within 1 year; ref. 44). We used any

date within the interval period based on the guidelines (i.e., 12-month period for mammography) when comparing self-report and medical record data for adherence to screening. We report FOBT and colorectal endoscopy (sigmoidoscopy or colonoscopy) separately to be comparable with other studies (1, 18, 26, 35). We combined sigmoidoscopy with colonoscopy as there were very few patients receiving sigmoidoscopy. For colorectal endoscopy, we excluded patients who were enrolled in the practice for less than 5 years because this could affect the accuracy of medical record documentation. We did not include barium enema in our reporting as there were no patients who received barium enema that did not have another form of colorectal cancer screening test. Patients who had a history of prior hysterectomy were excluded from the Pap smear screening outcome.

3. Behavioral counseling: Patient surveys and medical record reviews were also used to determine whether patients who did not ever have a cancer screening test had ever received a recommendation for that test. This question was not asked to patients who received the test, as we assumed that all patients who received the test received a recommendation (45-47). All patients were assessed for diet or weight counseling and exercise counseling within the past year. Patients who were current smokers were also assessed whether they received smoking cessation counseling within the past year.

We examined patient and practice characteristics as potential predictor variables of agreement. Patient characteristics were obtained from the patient survey and included age, gender, race/ethnicity, marital status, education, and health insurance. Practice characteristics of interest were obtained from the practice information survey and included specialty (family medicine, internal medicine, or mixed), years in existence, the number of clinicians, ratio of staff to clinician, and whether the practice had either nurse practitioners or physician assistants. Practice systems or resources related to cancer screening such as a system to communicate test results to patients, screening tools (cancer registry, screening checklist, or cancer referral system), the presence of electronic medical records, and onsite flexible sigmoidoscopy services were also obtained from the practice survey.

Statistical Analysis. For each risk factor, screening, or behavioral counseling question in the survey and item in the medical record review, binary variables were created for presence or absence of the documentation or patient report of the item. Number and percent of cases for each combination of self-report (yes/no) and chart documentation (yes/no) were calculated along with the overall percent of cases in which the two sources totally agreed (percent agreement on positives plus negatives). Because no recognized gold standard exists for these variables, sensitivities (percent true positives detected) and specificities (percent true negatives detected) were calculated two ways: (a) assuming the medical record as the gold standard and (b) assuming patient report as the gold standard. Hierarchical logistic regression models with random effects for practice were fit to calculate 95%

Table 1. Characteristics of sample in SCOPE, New Jersey, 2006-2007

Characteristic	Total n (%)
Patients	
Total sample	742 (100)
Age	
50-59	312 (42.0)
60-69	229 (30.9)
≥70	201 (27.1)
Gender	
Male	295 (39.8)
Female	447 (60.2)
Race*	
White	515 (69.8)
Black	128 (17.3)
Hispanic	61 (8.3)
Other	34 (4.6)
Marital status	
Married	470 (63.3)
Not married	272 (36.7)
Education level*	
Less than high school	88 (11.9)
High school diploma or some college	357 (48.4)
College or graduate school degree	292 (39.6)
Insurance*	
Private	343 (46.9)
Medicare	283 (38.7)
Medicaid	35 (4.8)
Other	38 (5.2)
None	33 (4.5)
Length of enrollment (y)	
≤1	152 (20.5)
2-4.9	261 (35.2)
5-9.9	229 (30.8)
≥10	100 (13.5)
Practices	
Total sample	25 (100)
Type of practice	
Family medicine	20 (80)
Internal medicine	3 (12)
Family and internal medicine	2 (8)
Years in business	
0-5	7 (28)
6-10	7 (28)
11-15	4 (16)
16-20	4 (16)
>20	3 (12)
Number of clinicians per practice	
1	3 (12)
2-4	13 (52)
5-7	7 (28)
≥8	2 (8)
Ratio of staff per clinician	
1-2	5 (20)
3-5	16 (64)
6-8	4 (16)
Midlevel providers	
None	15 (60)
Nurse practitioners	5 (20)
Physician assistants	3 (12)
Both nurse practitioner and physician assistant	2 (8)

*Numbers do not add to total due to missing data.

confidence intervals for the log-odds corresponding to each probability described above (agreement, sensitivity, or specificity). Estimation was conducted using pseudo-likelihood estimation (48). These intervals were then transformed to obtain the 95% confidence intervals for agreement, sensitivities, and specificities, adjusted for clustering of patients within practices.

We used hierarchical logistic regression to examine whether any patient or practice characteristic (from Table 1) contributed to differences in agreement between patient self-report and chart documentation. Nineteen patients were excluded from multivariate analysis due to missing covariate data. In particular, using chart documentation as a binary response (yes/no), analyses investigated whether patient or practice characteristics affected the explanatory effects of patient report. This was done using a score test of the interaction between each patient or practice characteristic and patient report based on generalized estimating equations with a working exchangeable correlation matrix (49). All analyses were done using SAS/STAT software (50).

Results

Patient and Practice Characteristics. There were 742 patients enrolled in the study with both patient survey and medical record review data. Table 1 summarizes the patient and practice characteristics. The majority of patients were between 50 and 70 years of age, white, female, currently married, had at least a high school education, and had private or Medicare health insurance. More than 55% of patients were enrolled in the practice for less than 5 years.

Most of the practices in our study were family medicine practices. Practices had been in existence an average of 11 years (SD 8.5). The mean number of clinicians per practice was 4.3 (SD 8.5), with a mean ratio of 3.4 staff per clinician (SD 1.4). Approximately 44% of practices ($n = 11$) had electronic medical records and 20% had onsite flexible sigmoidoscopy ($n = 5$).

Prevalence of Outcome by Data Source. Table 2 compares the prevalence of cancer-related risk factors, cancer screening, and behavioral counseling by patient report and medical record review. Patient report data showed higher prevalence rates for all outcomes, by 5% to 49%, except personal history of cancer, smoking status, and recommendations for mammography and Pap testing. Cancer screening rates were 12% (FOBT) to 35% (Pap test) higher by patient report than by medical record review.

Concordance, Sensitivity, and Specificity. Table 3 shows the measures of agreement for each outcome as well as the sensitivity and specificity taking either patient report or medical record review as the gold standard, after adjustment for clustering within practices. Rates of agreement were highest for risk factors and lowest for behavioral counseling and ranged from 41% (smoking cessation counseling) to 96% (personal history of cancer). Multivariate analyses showed that neither patient nor practice characteristics affected agreement rates.

Risk Factors. When assessing risk factors, sensitivity and specificity were similar whether medical records or patient surveys were used as the gold standard. The highest area of disagreement was when the patient reported a family history of cancer, but the medical record did not.

Cancer Screening. Agreement on cancer screening ranged from 61% (Pap and PSA testing) to 83%

Table 2. Prevalence of cancer-related risk factors, cancer screening, and behavioral counseling by data source among participants in SCOPE, New Jersey, 2006-2007

Item	Total <i>n</i>	Patient report	Medical record
		<i>n</i> (%)	<i>n</i> (%)
Cancer-related risk factors			
Personal history of cancer	742	63 (8.5)	71 (9.6)
Family history of cancer	742	161 (21.7)	125 (16.9)
Smoking status: current smokers	742	80 (10.8)	88 (11.9)
Cancer screening			
Mammography	447	310 (69.4)	161 (36.0)
Pap testing	319*	256 (80.3)	144 (45.1)
FOBT	742	180 (24.3)	92 (12.4)
Colorectal endoscopy	321 [†]	196 (61.1)	169 (52.7)
PSA testing	295	176 (59.7)	135 (45.8)
Behavioral counseling			
Recommendation for mammography [‡]	137	16 (11.7)	27 (19.7)
Recommendation for Pap testing [‡]	63	5 (7.9)	9 (14.3)
Recommendation for colorectal cancer screening [‡]	226	100 (44.3)	82 (36.3)
Diet or weight counseling in the past year	742	445 (60.0)	270 (36.4)
Exercise counseling in the past year	742	436 (58.8)	212 (28.6)
Smoking counseling for current smokers	80	66 (82.5)	27 (33.8)

NOTE: History of cancer includes breast, cervical, colorectal, or prostate cancer. Colorectal endoscopy includes sigmoidoscopy and colonoscopy procedures.

*One hundred twenty-eight patients were excluded due to prior hysterectomy.

[†]Four hundred twenty-one patients were excluded due to enrollment in practice for less than 5 y.

[‡]In patients who reported they did not ever receive the test.

(colorectal endoscopy). Sensitivity was high and specificity was low for all patient-reported tests except FOBT, when medical records were used as the gold standard, whereas the opposite was true when

patient surveys were used as the gold standard. Disagreements occurred most when the patient reported cancer screening, but the medical record did not have this documentation.

Table 3. Measures of agreement for risk factors, cancer screening, and behavioral counseling among participants in SCOPE, New Jersey, 2006-2007

Item	Medical record:	Yes	Yes	No	No	Agreement	MR as gold standard		PS as gold standard	
	Patient survey:	Yes	No	Yes	No		Sensitivity	Specificity	Sensitivity	Specificity
	Total <i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>		% (95% CI)*	% (95% CI)*	% (95% CI)*	% (95% CI)*
Cancer-related risk factors										
Personal history of cancer	723	52	19	11	641	96 (94-97)	73 (49-83)	98 (97-99)	83 (57-90)	97 (96-98)
Family history of cancer	723	71	50	85	517	81 (78-84)	59 (50-67)	86 (83-88)	46 (38-53)	91 (89-93)
Current smoker	723	64	22	14	623	95 (93-96)	74 (64-83)	98 (96-99)	82 (72-89)	97 (95-98)
Cancer screening										
Mammography	434	146	9	154	125	62 (58-67)	94 (89-97)	45 (34-58)	49 (43-54)	93 (87-98)
Pap testing	308	133	6	114	55	61 (53-67)	96 (91-98)	33 (26-40)	54 (43-60)	90 (80-96)
FOBT	723	46	44	130	503	76 (73-79)	51 [†]	79 (76-82)	26 [†]	92 (89-94)
Colorectal endoscopy	321	156	13	40	112	83 (79-87)	92 (87-96)	74 (66-80)	80 (73-85)	90 (82-98)
PSA testing	289	96	37	77	79	61 (51-67)	72 (64-79)	51 (43-58)	56 (39-63)	68 (59-76)
Behavioral counseling										
Recommendation for mammography [‡]	134	5	22	10	97	76 (68-83)	19 (0-28)	91 (84-95)	33 (0-56)	82 (74-88)
Recommendation for Pap testing [‡]	61	3	6	2	50	87 (45-90)	33 (0-99)	96 (86-99)	60 [†]	89 (49-93)
Recommendation for CRC screening [‡]	219	51	25	45	98	68 (62-74)	67 (58-87)	69 (57-76)	53 (43-63)	80 (72-86)
Diet or weight counseling	723	185	82	248	208	54 (51-58)	69 (64-75)	46 (41-50)	43 (38-47)	72 (67-82)
Exercise counseling	723	149	61	278	235	53 (49-57)	71 (65-77)	46 (42-50)	35 (31-40)	79 (74-84)
Smoking counseling	78	23	4	42	9	41 (19-52)	85 (34-94)	18 (9-31)	35 (11-46)	69 (41-88)
Formulas		<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	$(a+d)/(a+b+c+d)$	$a/(a+b)$	$d/(c+d)$	$a/(a+c)$	$d/(b+d)$

NOTE: History of cancer includes breast, cervical, colorectal or prostate cancer. Colorectal endoscopy includes sigmoidoscopy and colonoscopy procedures.

Abbreviations: MR, medical record; PS, patient survey; 95% CI, 95% confidence interval; CRC, colorectal cancer.

*95% Confidence intervals calculated based on pseudo-likelihood estimation for a hierarchical logistic regression model.

[†]95% Confidence intervals were not estimable due to small sample sizes or extreme clustering within practices.

[‡]In patients who reported they did not ever receive the test.

Behavioral Counseling. Agreement was moderate for recommendations for cancer screening, with higher specificity than sensitivity using either data source. The greatest source of disagreement for cancer screening recommendations was when patients reported receiving recommendations for colorectal cancer screening, but this was not documented in the medical record. In contrast, a recommendation for breast and cervical cancer screening was documented more frequently in medical records than reported by the patient. For diet or weight, exercise, and smoking cessation counseling, sensitivity was high and specificity was low for patient-reported counseling when medical records were used as the gold standard, whereas the opposite was true when patient surveys were used as the gold standard. Again, the greatest source of disagreement occurred when patients reported receiving counseling; however, it was not documented in the medical record.

Discussion

This is the first study to compare concordance between patient self-report and medical record data on cancer-related risk factors, cancer screening, and behavioral counseling among patients in community primary care practices. Agreement between patient report and medical record data was highest for risk factors (personal history of cancer and smoking status), moderate for cancer screening and recommendations for cancer screening, and lowest for behavioral counseling (smoking cessation, exercise, and diet or weight).

Our results showing high agreement between medical records and patient report on risk factors and low agreement on behavioral counseling are similar to previous findings (4, 28), but our agreement rates on cancer screening are lower than previously reported (11-13, 15, 17-20, 27, 31). Compared with weighted averages reported by Vernon et al., our sensitivity and specificity of self-report using the medical record as the gold standard are within the range reported by others for mammography, Pap testing, FOBT, and colorectal endoscopy (18). For PSA testing, our sensitivity of self-report is similar to others, whereas our specificity is considerably lower than others (18, 26), suggesting higher overreporting of these tests by our patient population. Because our patients were approached face-to-face before their doctor's visit, they might have overreported screenings due to social desirability influences (26). In addition, differences in patient populations and practice settings may explain some of these divergent findings because several prior studies were conducted in managed care settings, where medical record documentation may be more complete than in community primary care practices. Furthermore, the time interval used to define adherence to screening may have been different, such as using a 15- or 24-month period versus a 12-month period for mammography or FOBT, or using a 5-year versus a 10-year period when assessing colonoscopy (1, 18, 26, 32, 33, 51).

Similar to others, we found higher rates of cancer screenings reported by patients than documented in medical records, with patient report having high sensitivity but low specificity (18, 26). This difference between the two data sources could be due to patients

underestimating the time since their last screening test ("telescoping error"; refs. 12, 13, 34) or inaccurate reporting of screening due to recall errors, lack of knowledge, or social desirability bias (3, 15). In addition, patients may have received screenings outside of their primary care practice and this may have been inadequately tracked by these practices (1, 35). For example, some women may have directly gone to obstetrician/gynecologists for their well-women care, and these reports may not have been sent to their primary care physician, thus lowering our agreement rates for mammograms and Pap smears. However, we found similar agreement rates for prostate cancer screening, in which patients would usually not go elsewhere without referral. Even if testing is done outside of the primary care office, documentation of cancer screening status in the primary care medical record should be promoted as it is an important aspect of coordination of care activities essential to the patient-centered medical home (52).

These findings suggest that either method is good for assessing cancer risk factors. It is not surprising that there was high agreement between the two data sources on risk factors, because medical record documentation relies on the patient to give this information. Therefore, the data source for this information should be chosen based on the ease and expense of collecting the data. In most cases, this will mean that directly asking the patient for information on cancer risk factors is preferred.

The preferred data source for the other items will depend on the outcome of interest, the purpose, the resources available, and whether it is preferable to tolerate more false positives or false negatives. Medical record data is preferred for making clinical decisions to prescribe screening and when testing interventions because high false-positive rates from patient report may lead to delayed screening and potentially missed cancer diagnosis (26), as well as obscuring modest but potentially meaningful effects of an intervention (53). For example, Partin et al. showed that results of a mammography intervention trial differed significantly depending on whether administrative or patient self-report data were used. This difference in outcome results was due to bias from survey nonresponse (53). On the other hand, patient report may be preferable for assessing behavioral counseling, because medical record documentation is typically poor for services that are not specifically billable or that do not generate a report (7, 8). This lack of reimbursement for behavioral counseling may have decreased the accuracy of the medical record documentation. Health systems can address this deficiency in medical record documentation by adequate reimbursement of counseling services, whereas clinics can be encouraged to develop systems to better document behavioral counseling, such as using flow sheets. Self-report data may also be more feasible to obtain for population-based prevalence estimates of cancer screening; however, calculating correction factors to the self-report data or using multiple data sources is needed for more accurate public health surveillance (3, 26).

Almost 13% of male patients in this sample reported not receiving prostate cancer screening when a PSA test was documented in their medical record (28% false-negative rate). This is similar to the false-negative rates found by others (17, 19, 20). This is a cause of concern as

it suggests that physicians are ordering PSA tests as part of routine laboratory testing without discussion with the patient, contrary to strong recommendations for shared decision-making by the clinician and patient for prostate cancer screening (44).

There are several limitations to consider when interpreting the results of this study. First, it was conducted in 25 primary care practices in New Jersey with mostly white, married, educated, and insured patients; thus, the results may not be generalized to other populations and settings. Second, we did not have a gold standard to compare medical record or patient self-report data, and we did not validate patient report of cancer screenings with documentation from outside their primary care practice. Therefore, we cannot make judgments on the validity of either data source; we can only report how they compare with each other. Most studies determining validity of patient self-report data assume that the medical record is the gold standard and do not consider potential errors in documentation and abstraction. Third, our relatively small sample size and homogeneous population may not have allowed us to detect differences in agreement by patient or practice characteristics. For example, of 25 practices, our sample included only 5 practices that were either exclusively or partly internal medicine versus 20 that were exclusively family medicine practices. However, we did have enough power to detect differences due to patient characteristics for assessment of risk factors, cancer screening, and counseling for diet, weight, or exercise. Post hoc power calculations, based on comparisons of proportions, showed power >75%, in all but one case for detecting differences in the rate of agreement of 10% when comparing patients according to each patient-level characteristic listed in Table 1; in the remaining case, personal history of cancer versus race, this power was 67%. Fourth, the study sample was a convenience sample, which may have introduced some potential biases due to patient age, health, gender, employment status, time of year of recruitment, or other unmeasured factors. For example, we observed differences in the ages of patients who consented to participate in this study and the approximate ages of those who did not. Although this may have introduced some bias into our estimates of agreement, it is unclear whether a better response rate among younger patients would have increased or decreased concordance between the medical records and patient report. Most studies examining age as a predictor of agreement between medical records and patient self-report found age was not a significant factor (5, 11, 20, 54). Only one study found lower agreement for PSA tests and colorectal cancer screening with older age groups (17). Last, our study did not distinguish between screening tests and diagnostic tests and it is unclear how our results would have been affected if we did. Although Gordon et al. found high concordance between medical records and self-report about reasons for mammogram and Pap smear (25), others found low concordance on the reasons for mammogram (54), PSA, and colorectal cancer tests (17).

In conclusion, agreement between patient self-report and medical record data is highest for risk factors, moderate for cancer screening, and lowest for behavioral counseling among patients in community primary care practices. Implications of choosing self-report versus

medical records should be carefully considered when assessing health services for clinical practice, research, and public health surveillance. Deciding which data source to use will depend on the outcome of interest, whether the data is used for clinical decision making, performance tracking, or population surveillance; the availability of resources; and whether a false positive or a false negative is of more concern.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

References

- Schenck AP, Klabunde CN, Warren JL, et al. Data sources for measuring colorectal endoscopy use among Medicare enrollees. *Cancer Epidemiol Biomarkers Prev* 2007;16:2118–27.
- Beckles GL, Williamson DF, Brown AF, et al. Agreement between self-reports and medical records was only fair in a cross-sectional study of performance of annual eye examinations among adults with diabetes in managed care. *Med Care* 2007;45:876–83.
- Newell SA, Giris A, Sanson-Fisher RW, et al. The accuracy of self-reported health behaviors and risk factors relating to cancer and cardiovascular disease in the general population: a critical review. *Am J Prev Med* 1999;17:211–29.
- Tisnado DM, Adams JL, Liu H, et al. What is the concordance between the medical record and patient self-report as data sources for ambulatory care? *Med Care* 2006;44:132–40.
- Tisnado DM, Adams JL, Liu H, et al. Does concordance between data sources vary by medical organization type? *Am J Manag Care* 2007;13:289–96.
- Wu L, Ashton CM. Chart review. A need for reappraisal. *Eval Health Prof* 1997;20:146–63.
- Dresselhaus TR, Peabody JW, Lee M, et al. Measuring compliance with preventive care guidelines: standardized patients, clinical vignettes, and the medical record. *J Gen Intern Med* 2000;15:782–8.
- Stange KC, Zyzanski SJ, Smith TF, et al. How valid are medical records and patient questionnaires for physician profiling and health services research? A comparison with direct observation of patients visits. *Med Care* 1998;36:851–67.
- Chang ET, Smedby KE, Hjalgrim H, et al. Reliability of self-reported family history of cancer in a large case-control study of lymphoma. *J Natl Cancer Inst* 2006;98:61–8.
- Mitchell RJ, Brewster D, Campbell H, et al. Accuracy of reporting of family history of colorectal cancer. *Gut* 2004;53:291–5.
- Caplan LS, Mandelson MT, Anderson LA, et al. Validity of self-reported mammography: examining recall and covariates among older women in a Health Maintenance Organization. *Am J Epidemiol* 2003;157:267–72.
- Caplan LS, McQueen DV, Qualters JR, et al. Validity of women's self-reports of cancer screening test utilization in a managed care population. *Cancer Epidemiol Biomarkers Prev* 2003;12:1182–7.
- McGovern PG, Lurie N, Margolis KL, et al. Accuracy of self-report of mammography and Pap smear in a low-income urban population. *Am J Prev Med* 1998;14:201–8.
- Paskett ED, Tatum CM, Mack DW, et al. Validation of self-reported breast and cervical cancer screening tests among low-income minority women. *Cancer Epidemiol Biomarkers Prev* 1996;5:721–6.
- Pizarro J, Schneider TR, Salovey P. A source of error in self-reports of pap test utilization. *J Community Health* 2002;27:351–6.
- Mandelson MT, LaCroix AZ, Anderson LA, et al. Comparison of self-reported fecal occult blood testing with automated laboratory records among older women in a health maintenance organization. *Am J Epidemiol* 1999;150:617–21.
- Hall HJ, Van Den Eeden SK, Tolsma DD, et al. Testing for prostate and colorectal cancer: comparison of self-report and medical record audit. *Prev Med* 2004;39:27–35.
- Vernon SW, Briss PA, Tiro JA, et al. Some methodologic lessons learned from cancer screening research. *Cancer* 2004;101:1131–45.
- Jordan TR, Price JH, King KA, et al. The validity of male patients'

- self-reports regarding prostate cancer screening. *Prev Med* 1999;28:297–303.
20. Volk RJ, Cass AR. The accuracy of primary care patients' self-reports of prostate-specific antigen testing. *Am J Prev Med* 2002;22:56–8.
 21. Hoffmeister M, Chang-Claude J, Brenner H. Validity of self-reported endoscopies of the large bowel and implications for estimates of colorectal cancer risk. *Am J Epidemiol* 2007;166:130–6.
 22. Degnan D, Harris R, Ranney J, et al. Measuring the use of mammography: two methods compared. *Am J Public Health* 1992;82:1386–8.
 23. 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;147:556–62.
 24. Studts JL, Ghate SR, Gill JL, et al. Validity of self-reported smoking status among participants in a lung cancer screening trial. *Cancer Epidemiol Biomarkers Prev* 2006;15:1825–8.
 25. Gordon NP, Hiatt RA, Lampert DI. Concordance of self-reported data and medical record audit for six cancer screening procedures. *J Natl Cancer Inst* 1993;85:566–70.
 26. Rauscher GH, Johnson TP, Cho YI, et al. Accuracy of self-reported cancer-screening histories: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2008;17:748–57.
 27. Lawrence VA, De Moor C, Glenn ME. Systematic differences in validity of self-reported mammography behavior: a problem for intergroup comparisons? *Prev Med* 1999;29:577–80.
 28. Zhu K, McKnight B, Stergachis A, et al. Comparison of self-report data and medical records data: results from a case-control study on prostate cancer. *Int J Epidemiol* 1999;28:409–17.
 29. Bowman JA, Sanson-Fisher R, Redman S. The accuracy of self-reported Pap smear utilisation. *Soc Sci Med* 1997;44:969–76.
 30. Madlensky L, McLaughlin J, Goel V. A comparison of self-reported colorectal cancer screening with medical records. *Cancer Epidemiol Biomarkers Prev* 2003;12:656–9.
 31. Baier M, Calonge N, Cutter G, et al. Validity of self-reported colorectal cancer screening behavior. *Cancer Epidemiol Biomarkers Prev* 2000;9:229–32.
 32. Vernon SW, Tiro JA, Vojvodic RW, et al. Reliability and validity of a questionnaire to measure colorectal cancer screening behaviors: does mode of survey administration matter? *Cancer Epidemiol Biomarkers Prev* 2008;17:758–67.
 33. Partin MR, Grill J, Noorbaloochi S, et al. Validation of self-reported colorectal cancer screening behavior from a mixed-mode survey of veterans. *Cancer Epidemiol Biomarkers Prev* 2008;17:768–76.
 34. Champion VL, Menon U, McQuillen DH, et al. Validity of self-reported mammography in low-income African-American women. *Am J Prev Med* 1998;14:111–7.
 35. Schenck AP, Klabunde CN, Warren JL, et al. Evaluation of claims, medical records, and self-report for measuring fecal occult blood testing among Medicare enrollees in fee for service. *Cancer Epidemiol Biomarkers Prev* 2008;17:799–804.
 36. Navarro C, Chirlaque MD, Tormo MJ, et al. Validity of self reported diagnoses of cancer in a major Spanish prospective cohort study. *J Epidemiol Community Health* 2006;60:593–9.
 37. Manjer J, Merlo J, Berglund G. Validity of self-reported information on cancer: determinants of under- and over-reporting. *Eur J Epidemiol* 2004;19:239–47.
 38. 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;153:299–306.
 39. Kerber RA, Slattery ML. Comparison of self-reported and database-linked family history of cancer data in a case-control study. *Am J Epidemiol* 1997;146:244–8.
 40. Montaña DE, Phillips WR. Cancer screening by primary care physicians: a comparison of rates obtained from physician self-report, patient survey, and chart audit. *Am J Public Health* 1995;85:795–800.
 41. Crabtree BF, Miller WL, Stange KC. Understanding practice from the ground up. *J Fam Pract* 2001;50:881–7.
 42. Stroebel CK, McDaniel RR, Jr., Crabtree BF, et al. How complexity science can inform a reflective process for improvement in primary care practices. *Jt Comm J Qual Patient Saf* 2005;31:438–46.
 43. Centers for Disease Control and Prevention (CDC). National Health Interview Survey Questionnaire. Hyattsville (MD): National Center for Health Statistics; 2005.
 44. Smith RA, Cokkinides V, Eyre HJ. Cancer screening in the United States, 2007: a review of current guidelines, practices, and prospects. *CA Cancer J Clin* 2007;57:90–104.
 45. Ferrante JM, Chen PH, Crabtree BF, et al. Cancer screening in women: body mass index and adherence to physician recommendations. *Am J Prev Med* 2007;32:525–31.
 46. May DS, Kiefe CI, Funkhouser E, et al. Compliance with mammography guidelines: physician recommendation and patient adherence. *Prev Med* 1999;28:386–94.
 47. Wee CC, Phillips RS, McCarthy EP. BMI and cervical cancer screening among White, African-American, and Hispanic women in the United States. *Obes Res* 2005;13:1275–80.
 48. Wolfinger R, O'Connell M. Generalized linear mixed models: a pseudo-likelihood approach. *J Stat Comp Sim* 1993;4:233–43.
 49. Lipsitz SH, Laird NM, Harrington DP. Generalized estimating equations for correlated binary data: using the odds ratio as a measure of association. *Biometrika* 1991;78:153–60.
 50. SAS/STAT [computer program]. Version 9.1.3. Cary: SAS Institute, Inc.; 2004.
 51. Partin MR, Slater JS, Caplan L. Randomized controlled trial of a repeat mammography intervention: Effect of adherence definitions on results. *Prev Med* 2005;41:734.
 52. American Academy of Family Physicians, American Academy of Pediatrics, American College of Physicians, et al. Joint principles of the patient-centered medical home. March 2007. (Accessed November 19, 2007, at <http://www.medicalhomeinfo.org/Joint%20Statement.pdf>).
 53. Partin MR, Malone M, Winnett M, et al. The impact of survey nonresponse bias on conclusions drawn from a mammography intervention trial. *J Clin Epidemiol* 2003;56:867.
 54. Zapka JG, Bigelow C, Hurley T, et al. Mammography use among sociodemographically diverse women: the accuracy of self-report. *Am J Public Health* 1996;86:1016–21.

Self-report versus Medical Records for Assessing Cancer-Preventive Services Delivery

Jeanne M. Ferrante, Pamela Ohman-Strickland, Karissa A. Hahn, et al.

Cancer Epidemiol Biomarkers Prev 2008;17:2987-2994.

Updated version Access the most recent version of this article at:
<http://cebp.aacrjournals.org/content/17/11/2987>

Cited articles This article cites 51 articles, 12 of which you can access for free at:
<http://cebp.aacrjournals.org/content/17/11/2987.full#ref-list-1>

Citing articles This article has been cited by 7 HighWire-hosted articles. Access the articles at:
<http://cebp.aacrjournals.org/content/17/11/2987.full#related-urls>

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

**Reprints and
Subscriptions** To order reprints of this article or to subscribe to the journal, contact the AACR Publications
Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link
<http://cebp.aacrjournals.org/content/17/11/2987>.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's
(CCC)
Rightslink site.