Short Communication

Validity of Self-Reported Colorectal Cancer Screening Behavior

Monika Baier,[1] Ned Calonge, Gary Cutter, Maureen McClatchey, Sarah Schoentgen, Steve Hines, Al Marcus, and Dennis Ahnen

American Medical Center Cancer Research Center, Lakewood, Colorado 80214 [M. B., G. C., S. H., A. M.]; Kaiser-Permanente Health Care Program, Denver, Colorado 80231 [N. C.]; Kempe Prevention Research Center for Family and Child Health, Denver, Colorado 80218 [M. M.]; University of Colorado Health Sciences Center, Denver, Colorado 80262 [S. S., D. A.]; and Department of Veterans Affairs Medical Center, Denver, Colorado 80220 [D. A.]

Abstract

End points for trials promoting cancer screening are often based on self-reported screening behavior. This study was designed to evaluate and optimize the reliability of a computer-assisted telephone interview for collecting self-reported colorectal cancer screening behavior. Cases who had received a fecal occult blood test (FOBT), flexible sigmoidoscopy, and/or colonoscopy, and controls who had no record of colorectal screening were identified among 40–75-year-old members of the Denver Kaiser Permanente Health Care Program and were contacted by telephone. Sensitivities and specificities of self-reported screening were calculated by comparison of subjects' recall with Kaiser Permanente records. The questionnaire was revised based upon results of the pilot phase of the study. Using the revised questionnaire, the sensitivity of self-reported screening was 96.2% for the FOBT, 94.9% for flexible sigmoidoscopy, 88.7% for colonoscopy, and 96.2% for either endoscopic screening test. The specificity of self-reported screening was 85.9% for the FOBT, 92.2% for flexible sigmoidoscopy, 96.8% for colonoscopy, and 92.0% for either endoscopic screening test. No marked differences in the accuracy of the self-reports were detected as a function of gender, age, ethnicity, or family history of colorectal cancer of the participants. Self-reports of colon cancer screening behavior can be reliably used as end points for intervention trials when carefully phrased questions are used.

Introduction

Regular screening with FOBTs[2] and/or sigmoidoscopy has been shown to be effective in reducing mortality and morbidity from colorectal cancer (1–7). It is generally recommended that colorectal cancer screening with fecal occult blood testing, flexible sigmoidoscopy, or both should be offered to the average risk population over the age of 50 (5–7).

Because a family history of colorectal cancer is a risk factor for the disease (7, 8), it is generally recommended that screening begin by age 40 in first-degree relatives of patients with colorectal cancer, and colonoscopic screening is recommended if there is a strong family history of the disease.

Presently, colon cancer screening is performed in only 25–35% of the age-eligible population (9). Effective interventions to improve colon cancer screening rates could have a large impact on death from this disease. The end point of such intervention studies is completion of the requisite screening. Self-reports of colon cancer screening behavior, if valid and reliable, could greatly simplify such intervention trials, but there is little known about the accuracy of self-reported colorectal cancer screening behavior (10, 11).

We report the results of a validation study of self-report of colon cancer screening behaviors (FOBT, sigmoidoscopy, colonoscopy) in average risk participants and in those at high risk because of a family history of colorectal cancer.

Materials and Methods

Study Population

This study was approved by the Human Subject Committees of the University of Colorado Health Sciences Center and Kaiser Permanente Health Foundation. Men and women between the ages of 40 and 75 years who had been members of K-P continuously for at least 48 months as of April 1996 and hence, captured in their electronic database were eligible for participation.

Cases were defined as individuals who had returned FOBT cards to the clinic laboratory and/or had completed a flexible sigmoidoscopy, a colonoscopy, or both. Controls were participants who had no K-P record of any colorectal screening tests over the 4 prior years of K-P enrollment. The data were maintained in secured confidential files and used only for research purposes. Additionally, a list of patient identification numbers of a cohort of K-P members who reported having at least one first-degree relative with colorectal cancer was obtained. Computer-assisted telephone interviews were conducted among consenting subjects with and without first-degree relatives with colorectal cancer within both the case and control populations.

Questionnaire

The study was divided into two phases; the initial computer-assisted telephone interview questionnaire was tested in a pilot phase and modified for the final instrument.

Pilot Phase. The pilot phase was conducted in 46 controls (subjects with no recorded screening procedures in the previous 4 years) to evaluate the specificity of the responses to the questions. The controls were randomly chosen from the 29,000 controls supplied by K-P. Using the initial version of the questions, the specificity of the self-reports (the percent of

Received 10/15/99; revised 10/15/99; accepted 11/23/99.
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.

[1] To whom requests for reprints should be addressed, at AMC Cancer Research Center, Center for Research Methodology and Biometrics, 1600 Pierce Street, Lakewood, CO 80214. Phone: (303) 239-3491; Fax: (303) 239-3394; E-mail: baierm@amc.org.

[2] The abbreviations used are: FOBT, fecal occult blood test; K-P, Kaiser Permanente.

Table 1 Sensitivity for self-reports of colorectal cancer screening procedures

	FOBT		Flexible sigmoidoscopy		Colonoscopy		Colonoscopy or sigmoidoscopy	
	n	Sensitivity (%)	*n*	Sensitivity (%)	*n*	Sensitivity (%)	*n*	Sensitivity (%)
Overall	186	96.2 (93.6–99.0)[a]	79	94.9 (90.0–99.8)	97	88.7 (81.5–94.3)	159	96.2 (93.4–99.2)
Family history								
+FDR	29	100.0	10	90.0	18	94.4	25	100.0
−FDR	157	95.5	69	95.7	79	87.3	134	95.5
		(0.271)[b]		(0.418)		(0.373)		(0.263)
Gender								
Female	98	99.0	40	92.5	57	91.2	84	96.4
Male	88	93.2	39	97.4	40	85.0	75	96.0
		(0.054)		(0.322)		(0.298)		(0.889)
Ethnicity								
White	159	97.5	73	95.9	84	90.5	140	96.4
Non-white	27	88.9	6	83.3	13	76.9	19	94.7
		(0.032)		(0.162)		(0.131)		(0.734)
Age								
40–49	15	93.3	8	100.0	11	100.0	18	100.0
50–65	81	98.8	38	97.4	41	85.4	73	100.0
>65	90	94.4	33	90.9	45	88.9	68	91.2
		(0.242)		(0.568)		(0.542)		(0.011)

[a] 95% confidence intervals of the sensitivity.
[b] *P*s.

subjects who had not had the screening tests who correctly recalled not having had them) was 97.9% for colonoscopy and 85.4% for flexible sigmoidoscopy, but only 62.5% for FOBT. Debriefing interviews revealed that some participants considered samples of stools taken at the time of a rectal exam as an FOBT screening, and others had received FOBT screening tests at places (health fairs, federal employee programs) that were not captured in the Kaiser database. The description of an FOBT in the questionnaire was therefore revised to read, "A fecal occult blood test is when you are given a set of cards to take home and asked to smear a part of your stool on three separate occasions onto the cards and then to return the cards to be tested for blood." The subject was then asked, "Have you ever been given these cards to take home to test your stool for blood?" followed by, "Did you have this test at K-P?" and "Did you return the cards to K-P?" Debriefing interviews also indicated that patients had difficulty making a distinction between flexible sigmoidoscopy and colonoscopy based on the description given. Consequently, an additional qualifier was added to the description of colonoscopy: "Colonoscopy requires a preparation to clean the colon of stool, usually causing diarrhea, and the patient is given an i.v. with sedative medication."

Revised Questionnaire. The revised questionnaire was tested on a vanguard sample of 59 cases (participants who had had one or more of the screening tests in the previous 4 years). After completing the surveys for the 59 vanguard cases, no substantive changes in the interview were made so that the 59 vanguard cases were included in the final data analyses.

The final instrument was then used for the second phase of the validation study. The final questionnaire can be viewed in its entirety on our web site.[3] This validation phase consisted of calls to 100 controls who had no record of any of the screening tests and to 170 additional cases who had had one or more of the screening tests. These cases along with the vanguard cases gave a total sample of 229 cases. According to K-P records of the 229 cases, 186 had an FOBT (49 within the previous year), 79 had a sigmoidoscopy (15 within the previous year), and 97 had a colonoscopy (22 within the previous year).

Chart reviews were conducted for all patients whose self-report was in disagreement with the K-P database (false positives and false negatives).

Statistical Analysis

All analyses were performed using SAS version 6.10 for personal computers. Differences between proportions were assessed using a z statistic, whereas differences between groups on demographic variables were analyzed using a *t* test.

Results

Cases and controls were predominantly Caucasian (85.4% and 78.0%, respectively), with blacks representing 2.8% and 9% and Hispanics representing 6.8% and 7% of the sample, respectively. There was also a higher proportion of females (52.7% and 64.0%, respectively). Of the cases, 15.7% had a first-degree relative with cancer *versus* 5% in the controls. The subjects who had had one or more of the screening tests were older than those who had none of the tests (62.1 ± 8.4 years *versus* 54.4 ± 10.1 years; $P < 0.0001$).

The sensitivity of the colorectal cancer screening self-reports (Table 1) was calculated as the percentage of subjects correctly reporting having had the screening test among the cases who had the test according to the K-P records. Sensitivity for FOBT self-report was high (96.2%), and it was somewhat higher for flexible sigmoidoscopy (94.9%) than for colonoscopy (88.7%). If no differentiation was made between the two endoscopic screening procedures (either flexible sigmoidoscopy or colonoscopy), the sensitivity for self-reports of any endoscopic screening rose to 96.2%. To assess the consistency between the two methods of obtaining screening information, a κ statistic was computed. The κ for FOBT was 0.37, for flexible sigmoidoscopy 0.66, for colonoscopy 0.77, and for either endoscopic procedure 0.76. Females had a marginally significant higher recall of FOBT than did males (99% *versus* 93.2%; $P = 0.054$), but there were no gender differences for the other

[3] Internet address: www.amc.org/validation.htm.

Table 2 Specificity for self-reports of colorectal cancer screening procedures

	FOBT		Flexible sigmoidoscopy		Colonoscopy		Colonoscopy or sigmoidoscopy	
	n	Specificity (%)	*n*	Specificity (%)	*n*	Specificity (%)	*n*	Specificity (%)
Overall	64	85.9 (77.4–94.4)[a]	90	92.2 (86.7–97.7)	95	96.8 (93.3–100)	87	92.0 (86.3–97.7)
Family history								
+FDR	1	100.0	3	100.0	5	80.0	3	66.7
−FDR	63	85.7	87	92.0	90	97.8	84	92.9
		(0.271)[b]		(0.448)		(0.317)		(0.304)
Gender								
Female	38	84.2	58	98.3	62	98.4	56	98.2
Male	26	88.5	32	81.3	33	93.9	31	80.7
		(0.631)		(0.004)		(0.23)		(.004)
Ethnicity								
White	51	86.3	69	92.8	74	96.0	67	92.5
Non-white	13	84.6	21	90.5	21	100.0	20	90.0
		(0.881)		(0.928)		(0.364)		(0.712)
Age								
40–49	31	93.6	37	97.3	38	100.0	36	100.0
50–65	20	85.0	36	88.9	38	97.4	35	88.6
>65	13	69.2	17	88.2	19	89.5	16	81.3
		(0.087)		(0.347)		(0.101)		(0.018)

[a] 95% confidence intervals for specificity.
[b] *P*s.

screening tests (Table 1). Non-whites had a significantly poorer recall for having had a FOBT than did whites (97.5% *versus* 88.9%; $P = 0.032$), but it was not different for the other screening tests. There were also significant differences in sensitivity of reporting having had either of the two endoscopic screening tests as a function of age, with the oldest group (>65 years) having the poorest recall ($P = 0.011$). No differences in sensitivity for any of the tests were found between subjects as a function of whether the participant has a first-degree relative with colorectal cancer.

The accuracy of the self-reports for FOBT was not significantly different in the 41 participants who had had the test within the last year (97.6%) as compared to the 145 who had had the test longer than 1 year previously (95.9%).

The specificity of the self-reports (Table 2) was calculated as the percentage of subjects correctly reporting that they had not had the screening test among participants who had not had the test according to the K-P records. The overall specificity was 85.9% for FOBT, 92.2% for flexible sigmoidoscopy, 96.8% for colonoscopy, and 92.0% if no differentiation was made between the two endoscopic screening procedures. The κ statistics for the different screening procedures were 0.74, 0.34, 0.39, and 0.43, respectively. No difference in the specificity of the self-reports for any of the screening tests was found as a function of a family history of colorectal cancer or ethnicity. Women recalled more accurately than men that they had not had a sigmoidoscopy or either a sigmoidoscopy or colonoscopy (98.3% *versus* 81.3% and 98.2% *versus* 80.7%, respectively; $P = 0.004$). Specificity of recall when no differentiation was made between the two endoscopic procedures was significantly better in the younger than in the older age groups ($P = 0.018$).

Discussion

We found that after revision of our questionnaire following the pilot study, participants were able to provide highly accurate recall information about whether or not they had had any of the colon cancer screening tests. The sensitivity of self-reports for individuals who had had one or more of the tests within the previous 4 years ranged from 88.7 to 96.2% for the three screening tests. Similarly, subjects who had not had any of the three screening tests were highly accurate (85.9–96.8% specificity) in reporting not having had colorectal cancer screening within the previous 4 years.

The pilot phase of this study demonstrated the importance of precise wording of the questions. Modification of the description of a FOBT increased the specificity of the FOBT self-report from 64% in the pilot study to 85.9% with the final instrument. Even after modifying the pilot questionnaire to clarify the differences between flexible sigmoidoscopy and colonoscopy, subjects still had difficulty distinguishing between the two tests. The sensitivity of self-reports for flexible sigmoidoscopy (94.9%) or colonoscopy (88.7%) increased to 96.2% if no differentiation was made between the two procedures, indicating that subjects more accurately recall having had an endoscopic test than which test was performed. These results highlight the need to carefully word and then validate the questionnaire to obtain reliable self-reporting of colon cancer screening behavior. The results also show that even then a small false-positive rate will occur (3.8–11.3% in this study).

The accuracy of the self-reports of screening behavior in this study (sensitivities 88.7–96.2% and specificities 85.9–96.8%) are comparable to those for other cancer screening procedures and higher than the previous studies of colon cancer screening. Previous studies have reported that the sensitivity of self-reports of mammography screening ranged from 74 to 94% (12, 13, 14). There are few previous studies of the accuracy of self-reported colon cancer screening behavior (10, 11). Neither of the previous reports attempted to improve the accuracy of the instrument used to collect the self-reports. In the most extensive previous study of this issue, Gordon *et al.* (10) compared the self-reports from a mailed questionnaire to medical record audits and found sensitivities of 92% for the recall of FOBT but only 79% for the recall of sigmoidoscopy. The specificity of the self-reports was relatively low for both FOBT and sigmoidoscopy (70.6% and 87.5%). The low specificity results in the study by Gordon *et al.* (10) are similar to those we found in the pilot phase of our study (specificities of 62.5% for the recall of FOBT and 85.4% for sigmoidoscopy). By debriefing the par-

ticipants in the pilot study and revising the questionnaire, the specificity of self-reports for FOBT and sigmoidoscopy was increased to 85.9% and 92.2%, respectively, whereas high sensitivity (≥95%) was maintained. These results indicate that improved accuracy of self-reported colon cancer screening behavior can be obtained by carefully designing and pretesting the instrument used to obtain the self-reports.

Although the study by Lipkus *et al.* (11) was not primarily designed to estimate the validity of self-reported colorectal cancer screening, they did report a surprisingly low (<30%) specificity for self-reported FOBT screening among African-American users of a community health center. The authors concluded that investigators need to interpret self-reported data for colorectal cancer screening with caution. Our results lead to a very different conclusion. The overall specificities for colon cancer screening tests in our study (85.9–96.8%) were much higher than those reported by Lipkus *et al.* (11). Our data lead us to conclude that the results of a properly validated questionnaire of self-reported colon cancer screening behavior can be relied upon by investigators.

The reason(s) for the large discrepancy between our results and those reported by Lipkus *et al.* (11) are not clear, but it could be due to differences in the study populations, the instrument used to collect the data, or both. Our study participants were all adults enrolled in the K-P health plan. Although our study population was predominantly Caucasian (83%), we found only minor ethnicity-associated variability in the accuracy of the self-reports. The ranges of sensitivity and specificity for the self-reports were relatively high (76.9–94.7% and 84.6–100.0%, respectively) in the non-Caucasian participants in our study. The results raise the possibility that the discrepancy between the results of our study and those reported by Lipkus *et al.* (11) could be due to differences in socioeconomic status rather than to ethnicity alone. The sample of minorities in our study is too small, however, to exclude ethnicity-related differences in recall or reporting behavior. Our results suggest that the questionnaire that we developed may be suitable for Hispanic and black subjects in the K-P system, but this does not mean it would be valid in lower socioeconomic groups. We found that the sensitivity of self-reported colon cancer screening behavior was comparably high in subjects with and without a first-degree relative with colorectal cancer (90–100% *versus* 87.3–95.7%, respectively). This observation indicates that self-reported screening behavior may be reliably used both in the average risk population and in populations at increased colon cancer risk due to a positive family history.

We could find no previous studies of the accuracy of self-reports of colonoscopy screening. Not surprisingly, we found similarly high sensitivity and specificity of self-reports for colonoscopy (88.7% and 96.8%, respectively) as we did for sigmoidoscopy (94.9% and 92.2%, respectively), but as noted above, participants had some residual difficulty distinguishing between the two types of endoscopic screening tests.

We did not identify any subgroup of participants that had markedly inaccurate self-reports. The overall accuracy of self-reports for colon cancer screening tests may have declined in older participants in our study. Although the differences were only statistically different for the sensitivity and specificity of self-reports for either of the two endoscopic tests, the relative consistency of the age effect among all of the screening tests suggests that age-related decline in self-report accuracy may be real rather than a chance finding. Nonetheless, even the over-65-year-old participants had a reasonably high accuracy of the self-reports.

There are several limitations in our study design. It is possible that the population in the K-P system is different from other groups. Kaiser members are regularly surveyed about their medical care and might be more willing to respond to this survey than the general population. By definition, the K-P group had health insurance, which means that the results of this study may not be applicable to subjects who do not have health insurance. The large majority of participants in our study were white (83%). We did find a statistically lower sensitivity for FOBT self-reports in the non-white group (88.9 *versus* 97.5%; $P = 0.032$), but we found no other differences in accuracy between whites and non-whites. Nonetheless, our results may not be generalizable to minority populations because of the small representation of minorities in our study population. Finally, the subset analyses of this data set by age and family history of colorectal cancer should be considered exploratory because of the relatively small sample sizes and the multiple comparisons made.

Despite its limitations, this study establishes that the questionnaire that we developed can provide valid self-reported colon cancer screening information in the population we studied. We believe it to be suitable for use in behavioral intervention protocols in similar populations.

References

1. Mandel, J. S., Bond, J. H., Church, T. R., Snover, D. C., Bradley, G. M., Schuman, L. M., and Ederer, F. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. N. Engl. J. Med., *328:* 1365–1371, 1993.

2. Kronborg, O., Fenger, C., Olsen, J., Jorgensen, O. D., and Sondergaard, O. Randomised study of screening for colorectal cancer with faecal occult blood test. Lancet, *348:* 1467–1471, 1996.

3. Hardcastle, J. D., Chamberlain, J. O., Robinson, M. H., Moss, S. M., Amar, S. S., Balfour, T. W., James, P. D., and Mangham, C. M. Randomized controlled trial of faecal occult blood screening for colorectal cancer. Lancet. *348:* 1472–1477, 1996.

4. Selby J. V., Friedman, G. D., Quesenberry, C. P., and Weiss, N. S. A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. N. Engl. J. Med., *326:* 653–657, 1992.

5. Woolfe, S. H. Screening for Colorectal Cancer. Guide to Clinical Preventive Services. Report of the U. S. Preventive Services Task Force, pp. 89–104. Baltimore, MD: Williams & Wilkins, 1996.

6. Byers, T., Levin, B., Rothenberger, D., Dodd, G. D., and Smith, R. A. American Cancer Society guidelines for screening and surveillance for early detection of colorectal polyps and cancer: update 1997. American Cancer Society Detection and Treatment Advisory Group on Colorectal Cancer. CA Cancer J. Clin., *47:* 154–160, 1997.

7. Winawer, S. J., Fletcher, R. H., Miller, L., Godlee, F., Stolar, M. H., Mulrow, C. D., Woolf, S. H., Glick, S. N., Ganiats, T. G., Bond, J. H., Rosen, L., Zapka, J. G., Olsen, S. J., Giardiello, F. M., Sisk, J. E., Van Antwerp, R., Brown-Davis, C., Marciniak, D. A., and Mayer, R. J. Colorectal cancer screening: clinical guidelines and rationale. Gastroenterology, *112:* 594–642, 1997.

8. Fuchs, C. S., Giovannucci, E. L., Colditz, G. A., Hunter, D. J., Speizer, F. E., and Willet, W. C. A prospective study of family history and the risk of colorectal cancer. N. Engl. J. Med., *331:* 1669–1674, 1994.

9. Vernon, S. W. Participation in colorectal cancer screening. J. Natl. Cancer Inst., *89:* 1406–1422, 1997.

10. Gordon, N. P., Hiatt, R. A., and Lampert, D. I. Concordance of self-reported data and medical record audit for six cancer screening procedures. J. Natl. Cancer Inst., *85:* 566–570, 1993.

11. Lipkus, I. M., Rimer, B. K., Lyna, P. R., Pradhan, A. A., Conaway, M., and Woods-Powell, C. T. Colorectal screening patterns and perceptions of risk among African-American users of a community health center. J. Community Health, *21:* 409–427, 1996.

12. Paskett, E. D., Tatum, C. M., Mack, D. W., Hoen, H., Case, L. D., and Velez, R. Validation of self-reported breast and cervical cancer screening tests among low-income minority women. Cancer Epidemiol. Biomark. Prev., *5:* 721–726, 1996.

13. Degnan, D., Harris, R., Ranney, J., Quade, D., Earp, J. A., and Gonzalez, J. Measuring the use of mammography. Two methods compared. Am. J. Public Health, *82:* 1386–1388, 1992.

14. King, E. S., Rimer, B. K., Trock, B., Balshem, A., and Engstrom, P. How valid are mammography self-reports? Am. J. Public Health, *80:* 1386–1388, 1990.