Fall 2014

Colorectal cancer screening in the Iowa Research Network (IRENE): a validity assessment of patient self-report of up-to-date status

Carol Ann Moss
University of Iowa

COLORECTAL CANCER SCREENING IN THE IOWA RESEARCH NETWORK (IRENE): A VALIDITY ASSESSMENT OF PATIENT SELF-REPORT OF UP-TO-DATE STATUS

by

Carol Ann Moss

A thesis submitted in partial fulfillment
of the requirements for the
Master of Science degree in Epidemiology
in the Graduate College of
The University of Iowa

December 2014

Thesis Supervisors: Professor Charles Lynch
Professor Barcey T. Levy

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

MASTER'S THESIS

_____

This is to certify that the Master's thesis of

Carol Ann Moss

has been approved by the Examining Committee
for the thesis requirement for the Master of Science degree
in Epidemiology at the December 2014 graduation.

Thesis Committee: _____
Charles Lynch, Thesis Supervisor

_____
Barcey T. Levy, Thesis Supervisor

_____
Trudy L. Burns

_____
Ryan Carnahan

ACKNOWLEDGEMENTS

The author thanks Barcey T. Levy, PhD, MD, whose funded project and wonderings spurred the origination of this thesis, Trudy L. Burns, MPH, PhD, for her steadfast support with regard to the intricacies of data analysis, Charles Lynch, MD, MS, PhD, for his mentorship and overall guidance on matters well beyond this thesis, and Ryan Carnahan, PharmD, MS, for his astute insights and reliably "big picture" contributions. All gave of themselves with repeated unfailingly-critical readings.

The author gratefully acknowledges Trudy L. Burns, MPH, PhD, for use of her "Statistical Methods in Epidemiology Fall 2012" course handouts and Anne Wallis, MHS, PhD, for her responses to the initial presentation of some of the material contained in this thesis.

ABSTRACT

Patient self-report of colorectal cancer (CRC) screening remains a critical source of information in determining adherence to recommended guidelines. Accurate assessment is important for clinical decision-making, quality assurance and research. Population subgroup differences can affect self-report accuracy. Studies relying on patient self-report benefit from assessing validity and attempting to quantify measurement error and bias. This study assessed self-reported CRC screening data accuracy – estimating overall and test-specific adherence using four common validity measures (sensitivity, specificity, concordance, and report-to-records ratio [R2R]) – and evaluated associations between predictor variables and accuracy that might explain variation in estimates.

1,399 patients aged 51-80 years from 16 family medicine offices in the Iowa Research Network (IRENE), a practice-based research network (PBRN), completed an investigator-developed questionnaire and had medical records (MRs) available. Comparison of self-report of up-to-date screening with test documentation in the MR was used to estimate validity; multivariable analysis assessed predictors of concordance, or agreement between self-report and test documentation in the MR, for colonoscopy (CSPY). Predictor variables included

patient characteristics (age, gender, education, income, insurance status, family history of CRC and IRENE office), healthcare utilization practices (recency of last visit to office and duration of patient status in office) and patient rural-urban residence classified according to a four-category Rural-Urban Commuting Area (RUCA) coding scheme.

Sixty percent of patients reported they were up-to-date with CRC screening by any test, while 48% had screening documented in the MR (sensitivity 0.95, specificity 0.73, concordance 0.83 and R2R 1.24). Nearly all documentation was for CSPY (sensitivity 0.94, specificity 0.76, concordance 0.84 and R2R 1.21). Education, insurance source, CRC family history and patient duration in office, when adjusted for all other variables in the final model, were significant ($p < 0.05$) predictors of concordance. Age modified a significant association with concordance for patient rural-urban residence ($p = 0.03$) and for recency of last visit ($p = 0.04$).

Self-reported CRC screening validity was generally acceptable, but overreporting was prevalent across all tests. MR documentation of CRC screening was almost exclusively based on CSPY. Concordance between self-reported CSPY and the MR was good but varied with patient characteristics, healthcare utilization practices and rural-urban residence.

PUBLIC ABSTRACT

Patient self-report of colorectal cancer (CRC) screening remains a critical source of information in determining adherence to recommended guidelines. Accurate assessment is important for clinical decision-making, quality assurance and research. Population subgroup differences can affect self-report accuracy. Studies relying on patient self-report benefit from assessing validity and attempting to quantify measurement error and bias. This study assessed self-reported CRC screening data accuracy, estimating overall and test-specific adherence, and evaluated associations between predictor variables and accuracy that might explain variation in estimates.

1,399 patients aged 51-80 years from 16 family medicine offices completed an investigator-developed questionnaire and had medical records (MRs) available. Comparison of self-report of up-to-date screening with test documentation in the MR was used to estimate validity; multivariable analysis assessed predictors of concordance, or agreement between self-report and test documentation in the MR.

Sixty percent of patients reported they were up-to-date with CRC screening by any test, while 48% had screening documented in the MR. Nearly all documentation was for colonoscopy. Education, insurance source, CRC

family history and patient duration in office, adjusted for all other variables in the final model, were significant predictors of concordance. Age modified a significant association with concordance for patient rural-urban residence and for recency of last visit.

Self-reported CRC screening validity was generally acceptable, but overreporting was prevalent across all tests. MR documentation of CRC screening was almost exclusively based on colonoscopy. Concordance between self-reported colonoscopy and the MR was good but varied with patient characteristics, healthcare utilization practices and rural-urban residence.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACS	American Cancer Society
AGA	American Gastroenterological Society
AOR	Adjusted odds ratio
BRFSS	Behavioral Risk Factor Surveillance System
CDC	Centers for Disease Control and Prevention
CRC	Colorectal Cancer
CSPY	Colonoscopy
CI	Confidence interval
DCBE	Double-contrast Barium Enema
EMR	Electronic medical record
FIT	Fecal Immunochemical Test
FN	False negative
FOBT	Fecal Occult Blood Test
FP	False positive
FSIG	Flexible Sigmoidoscopy
GED	General Education Diploma
HINTS	Health Information National Trends Survey
HIPAA	Health Information Portability and Accountability Act
HMO	Health Maintenance Organization
HRSA	Health Resources and Service Administration
IRENE	Iowa Research Network
MR	Medical record
NCI	National Cancer Institute
NHIS	National Health Interview Survey
OR	Odds ratio
PBRN	Practice-based Research Network
RUCA	Rural-Urban Commuting Area
R2R	Report-to-records ratio
SD	Standard deviation
TN	True negative
TP	True positive
USPSTF	United States Preventive Services Task Force
UTD	Up-to-date
VA	Veterans Affairs
WWAMI	Washington, Wyoming, Alaska, Montana and Idaho

INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer among males and females in the United States (U.S.) and the second leading cause of cancer death (American Cancer Society, 2014). The same ranking holds for the incidence and mortality rates for residents of the state of Iowa (State Health Registry of Iowa, 2014).

Evidence suggests that early detection and removal of adenomas reduces morbidity and mortality from CRC (Zauber et al, 2012). National scientific panels, organizations and advocacy groups, including the U.S. Preventive Services Task Force (USPSTF), the American Gastroenterological Association (AGA) and the American Cancer Society (ACS), issue and update recommendations for CRC screening for average risk individuals on a recurring basis. Accurate assessment of adherence to these guidelines is important for clinical decision making, quality assurance and research (Partin et al, 2008).

Data sources for estimates of CRC screening prevalence include claims data, medical records and patient self-report. Population-based estimates of the prevalence of CRC screening in the U.S. have relied primarily on patient self-reports through continuing household data collection methods, including the Centers for Disease Control and Prevention's (CDC) in-person National Health Interview Survey (NHIS) and telephone Behavioral Risk Factor Surveillance System (BRFSS) (Centers for Disease Control and Prevention, retrieved March 30, 2014) and the National Cancer Institute's (NCI) Health Information National Trends Survey (HINTS), a relatively recent addition to U.S. population-level surveys, unique in its emphasis on cancer, health communication and health information technology (National Cancer Institute, retrieved April 5, 2014).

Obtaining reliable and valid measures of screening behaviors has proven challenging (Zapka, 2008). Reiter et al (2013) note that accurate recall of CRC

screening within guidelines may be even more difficult compared with screening for other cancers (e.g., breast, cervical and prostate) because of the larger number of available and recommended CRC screening tests. These tests, which include fecal occult blood tests (FOBT), fecal immunochemical tests (FIT), flexible sigmoidoscopy (FSIG), double contrast barium enema (DCBE) and colonoscopy (CSPY), are recommended at differing intervals, the recommendations change over time, and some patients have limited familiarity with the various tests (Shokar et al, 2011).

Self-report remains a critical source of information on CRC screening adherence for several reasons, including, most obviously, that this is often the *only* information source (Partin et al, 2008). In addition, self-report is the most efficient means for collecting screening adherence information, especially since the various CRC screening procedures may not all be conducted in the same clinical settings and thus not regularly documented in the medical record at a single location. As patients may receive their medical care across several offices over time, using medical records to assess accuracy is further limited. Even if an individual obtains care within a single location or health system, specific results of CRC screening tests are filed in multiple areas of the medical record, whether paper or electronic, making chart reviews challenging. (FOBT/FIT, for example, may be filed under "Labs", CSPY under "Hospital Procedures" and DCBE under "Imaging.") Moreover, implementation of federal legislation limiting access to medical records, i.e., the Health Information Portability and Accountability Act (HIPAA), complicates the gathering of medical data from more than one source and increases the expense, thus making self-report an efficient way to obtain CRC screening data.

Data from validation studies are important for both practice and policy decisions (Zapka, 2008). Clinical implications of adherence by patients are

exemplified by those who have more than one source of medical care, as they may not receive timely CRC screening if providers base their recommendation solely on potentially-inaccurate patient self-reports. Zapka (2008) concludes that, although adjustment can be made to national screening prevalence estimates, overreporting will continue to challenge practitioners' ability to determine the accuracy of individual patient reports.

Authors of previous CRC screening studies have emphasized that studies using self-reported screening adherence should assess reliability and validity and attempt to quantify measurement error and bias (Meissner, Vernon et al, 2004; Vernon, Meissner et al, 2004). As they suggest, self-reported data can be misleading if they are systematically biased. Substantial variation across studies in terms of how self-reported CRC screening adherence data are collected and reported can make it difficult to draw conclusions (Partin et al, 2008). Thus, the increased need for and subsequent use of self-report as the main data source for CRC screening adherence points to the importance of validity studies (Zapka, 2008). It is with this in mind that the present validity study was undertaken.

REVIEW OF THE LITERATURE

Articles published in Medline between January 1, 2000, and February 28, 2014, were searched. The year 2000 was chosen as the earliest year to minimize the inclusion of digital rectal exams, an earlier common screening method but one which by 2002 was no longer considered acceptable by the USPSTF (U.S. Preventive Services Task Force, 2008). Article titles containing the words "colorectal cancer screening," "self-report," and "validity" were sought. The references in the identified articles were also used to locate additional articles of interest. Lastly, related citations for each article identified by the previous methods, as listed on Medline, were reviewed.

Twenty-four articles were identified presenting data on self-reported and medical record- or claims-documented CRC screening histories. Six studies were excluded: one that examined the validity of self-reported CRC screening *intentions* rather than actual screening behavior (Lipkus et al, 2000), a second that reported validity estimates for combined CSPY and FSIG only (Khoja et al, 2007), a third (Hoffmeister et al, 2007) that involved *CRC cases* and their matched controls as participants, another that dealt with the reliability and validity of self-reported *diagnosis* of CRC rather than CRC screening (Lynch et al, 2008), and a fifth that was a meta-analysis of studies published before the initial cutoff date of January 1, 2000 (Rauscher et al, 2008). A more recently-published sixth article (Vernon et al, 2012) re-examined data from a previously-published and included study (Vernon et al, 2008) but did not present any new validity estimates and thus was excluded.

These six exclusions left 18 published articles on the accuracy of self-reported CRC screening (Baier et al, 2000; Bastani et al, 2008; Ferrante et al, 2008; Fiscella et al, 2006; Griffin et al, 2009; Hall et al, 2004; Jones et al, 2008; Lipkus et al, 2003; Madlensky et al, 2003; Partin et al, 2008; Powe et al, 2008; Schenck et al,

2008; Schoen et al, 2002; Shokar et al, 2011; Vernon et al, 2008; White et al, 2013; Reiter et al, 2013; Beebe et al, 2014). One study (Madlensky et al, 2003) was conducted outside of the U.S. All included studies compared self-reported information on CRC screening to medical records only, except for Fiscella et al (2006), which compared self-reported information with Medicare claims data, and Schenck et al (2008), which used both medical records and Medicare claims data in comparison with self-reports.

Sixteen of the 18 included studies assessed validity of the *entire* study sample according to at least one of the common validity estimates, defined below. These studies are summarized in Table A1 and are referred to as "Primary Validity Analysis Studies." Two additional studies (Griffin et al, 2009; White et al, 2013) were secondary analyses of the 16 studies presented in Table A1. These, along with seven of the studies in Table A2, present validity estimates stratified by population subgroup and/or identify correlates of accurate self-reported screening using logistic regression analysis. The results of these nine studies are summarized separately in Table A2, "Secondary Validity Analysis Studies."

Primary Validity Analysis Studies

Details of each Primary Validity Analysis study (n = 16) were summarized, including the age range of participants, the setting from which they were selected, including geographic locale, and the sociodemographic characteristics of the sample, including race, gender and socioeconomic status, if specified (see Table A1). In addition, the recency of specific tests was recorded and included "ever" or "up-to-date", based on USPSTF recommendations for specific CRC screening tests. For studies that calculated measures of recency, only the most recent measure was included in the summary.

Measures of Validity

Because accuracy of self-report is a multidimensional concept, i.e., screening test information can be reported for *any* test in a given time period or the most recent test, no single measure of accuracy is sufficient (Sudman et al, 1994). Vernon et al (2008) recommend that numerous analytic methods be used to assess validity of self-report of CRC screening. Four common measures examining validity have been called for by Zapka to encourage similar reporting across studies (Zapka, 2008). These measures, which include sensitivity, specificity, concordance and report-to-records ratio (R2R), were noted, when possible, for overall adherence by any test and for test-specific adherence. They are defined as follows:

Sensitivity: proportion of participants who had documentation of a test in the medical record who reported having had the test, or TP/(TP + FN), where TP = true positive and FN = false negative.

Specificity: proportion of participants without documentation of a test in the medical record who reported not having had the test, or TN/(TN + FP), where TN = true negative and FP = false positive.

Concordance: agreement between self-report and medical record/claims data, or (TP + TN)/total sample, and

Report-to-records ratio (R2R): total number of self-reported tests relative to actual number documented, or (TP + FP)/(TP + FN).

R2R, a measure of net bias, can also be viewed as the number of patients with positive self-reports divided by the number of patients who actually received the test according to the record source (Vernon, Meissner et al, 2004). Interpretation of the R2R is straightforward: R2R > 1.00 indicates overreporting; R2R < 1.00 indicates underreporting. However, according to Partin et al (2008),

historically, there has been no consensus regarding acceptable ranges for the other three validity measures in the context of CRC screening.

Vernon et al (2008) suggested that acceptable values for these estimates will vary depending on how they are used. For example, if the purpose is to evaluate the efficacy of a behavior change intervention, a measure with low sensitivity may be acceptable. If the purpose is to identify individuals who have been screened for CRC, as is the case of the current thesis, then it would be desirable to have high sensitivity and accept a lower specificity in order not to miss those who remain unscreened. Several authors (Partin et al, 2008; Vernon et al, 2008; Shokar et al, 2011; White et al, 2013; Reiter et al, 2013) of the studies reviewed used the criteria of Tisnado et al (2006) to evaluate sensitivity, specificity and concordance: an estimate < 0.7 indicates poor agreement; ≥ 0.7 to < 0.8 indicates fair agreement; ≥ 0.8 to < 0.9 indicates good agreement; and ≥ 0.9 indicates excellent agreement.

Results of Analysis of the Primary Validity Literature

For details of the Primary Validity Analysis studies, see Table A1. The two main findings from this primary analysis of the literature include: 1) validity of self-report, as measured by estimates of sensitivity, specificity and concordance, was generally respectable (≥ 0.7) across studies but varied by particular screening test; and 2) overreporting was found in nearly all studies in which an estimate of R2R was provided. Similar conclusions were reached in the meta-analysis of Rauscher et al (2008).

For CRC screening by any test, sensitivity across the studies that included such estimates (six of 16) ranged from 0.85 (Shokar et al, 2011) to 0.98 (Partin et al, 2008); specificity (six of the 16 studies) from 0.58 (Reiter et al, 2013) to 0.81 (Jones et al, 2008); concordance (five of the 16 studies) from 0.74 (Bastani et al,

2008) to 0.89 (Beebe et al, 2014); and R2R (six of the 16 studies) from 1.03 (Beebe et al, 2014) to 1.62 (Bastani et al, 2008).

For FOBT, sensitivity (13 of 16 studies) ranged from 0.32 (Reiter et al, 2013) to 0.96 (Baier et al, 2000); specificity (13 of 16 studies) from 0.50 (Powe et al, 2008) to 0.93 (Reiter et al, 2013); concordance (11 of 16 studies) from 0.50 (Powe et al, 2008) to 0.90 (Reiter et al, 2013); and R2R (12 of 16 studies) from 0.76 (Lipkus et al, 2003) to 3.14 (Madlensky et al, 2003).

For CSPY, sensitivity (nine of 16 studies) ranged from 0.84 (Bastani et al, 2008) to 1.00 (Jones et al, 2008); specificity (nine of 16 studies) from 0.65 (Reiter et al, 2013) to 0.97 (Baier et al, 2000); concordance (seven of 16 studies) from 0.80 (Reiter et al, 2013) to 0.94 (Madlensky et al, 2003); and R2R (eight of 16 studies) from 0.92 (Beebe et al, 2014) to 1.52 (Bastani et al, 2008).

For FSIG, sensitivity (nine of 16 studies) ranged from 0.17 (Reiter et al, 2013) to 0.95 (Baier et al, 2000); specificity (nine of 16 studies) from 0.76 (Partin et al, 2008) to 0.97 (Bastani et al, 2008); concordance (seven of 16 studies) from 0.76 (Partin et al, 2008; Beebe et al, 2014) to 0.96 (Bastani et al, 2008; Reiter et al, 2013); and R2R (eight of 16 studies) from 0.80 (Beebe et al, 2014) to 4.50 (Reiter et al, 2013).

For DCBE, which was less frequently reported, sensitivity (five of 16 studies) ranged from 0.45 (Shokar et al, 2011) to 0.92 (Beebe et al, 2014); specificity (five of 16 studies) from 0.72 (Beebe et al, 2014) to 0.97 (Vernon et al, 2008); concordance (four of 16 studies) from 0.72 (Beebe et al, 2014) to 0.92 (Vernon et al, 2008); and R2R (four of 16 studies) from 0.82 (Vernon et al, 2008) to 22.30 (Beebe et al, 2014).

Sources of Variability in Validity Studies

Sources of variability in validity studies can be divided into roughly four components: those related to the 1) survey instrument, 2) medical records, 3)

analysis or 4) characteristics of the patient (Vernon, Meissner et al, 2004; Partin et al, 2008, Shokar et al, 2011).

Instrument-related Variability

Cognitive issues

Survey design features affect accuracy of self-reports (Zapka, 2008). Sudman et al (1994) noted that participants in CRC screening adherence studies are asked to recall screening histories across a variety of time intervals, ranging from immediately after a medical encounter up to the previous five or, now with the emphasis on CSPY, even 10 years. Vernon, Meissner et al (2004) suggest that when tests are similar to each other, misreporting may occur. Patients may not be aware which test was administered, e.g., FSIG and CSPY, or understand the differences between tests, such as upper and lower endoscopies.

In the study by Partin et al (2008), for example, based on data from an earlier time, overreporting was most pronounced for CSPY. These authors explain that this may have reflected the fact that this method of screening was the newest of the four CRC screening procedures at the time and thus was the one with which patients were least familiar. In a more recent study, Shokar et al (2011) report secular trends in the U.S. in the use of CRC tests that point to greater use of CSPY and declining rates of FOBT, FSIG and DCBE (Meissner et al, 2006; Rim et al, 2011). They conclude that the prevalence of inaccurate validity estimates may decline as CSPY becomes the most prevalent CRC screening test.

Development of survey items in this context can be difficult. An individual's responses to questions are a complex process that has been conceptualized as involving four different stages, each of which can lead to inaccuracies in reporting (Sudman et al, 1994; Shokar et al, 2011).

During the first stage, labeled "Comprehension" by Shokar et al (2011), the individual must interpret the meaning of the question. Questions detailing

specific CRC screening procedures may be subject to linguistic and interpretive difficulties (Sudman et al, 1994). Not all participants, especially those with less education, for example, may understand what a fecal occult blood test is. This may lead the participant to incorrectly report either that a test occurred when it did not, or that a test did not occur when it did. The solution to this problem, according to the authors, is to choose the most meaningful wording of questions, including alternative word choices for different populations of respondents.

The second stage, "Retrieval", involves retrieving either the answer to the question or relevant information that can be used to construct an answer, and, therefore, reliance on long-term memory (Sudman et al, 1994). Here the wording of a question becomes particularly important as it acts as a cue to improve retrieval.

In the third stage, "Estimation/Judgment", the individual assesses the information retrieved and its relevance to the question and makes a choice to accept or reject the information. Judgments are again based on data retrieved from memory.

Finally, in the fourth or "Response" stage, the individual weighs factors such as social desirability and may choose to edit a response if certain answers are seen as being more socially desirable than others. Socially desirable information is overreported whereas undesirable behavior is underreported (Sudman et al, 1994). According to these authors, CRC screening tests may be seen by female respondents as something that is socially desirable, and so they may be reluctant to report *not* getting these tests. Thus overreporting of screening tests could occur. Vernon et al (2012) investigated the effect of social desirability on the accuracy of self-reported CRC screening tests in a recent paper but found no significant association.

Telescoping

A related cognitive aspect of self-report is that of "telescoping." Telescoping is a well-recognized problem in recalling when an event occurred: the event is remembered, but *when* it happened is misremembered (Sudman et al, 1994). It is a form of recall bias regarding the time period and is another possible explanation for overreporting of CRC screening. Forward telescoping, or remembering that an event occurred more recently than it actually did, is more common than backward telescoping (Sudman et al, 1994).

Taking this concept into account, Vernon et al (2008) concluded from an ancillary analysis of their study data that the framing of CRC screening questions regarding time should be in terms of a specified interval rather than open-ended. They call for future research that identifies the cognitive as well as the motivational barriers and facilitators for accurately self-reporting CRC screening adherence to ensure that data are both reliable and valid estimates of screening behavior.

Survey administration mode

Vernon, Meissner et al (2004) suggest that survey mode (telephone, mail, face-to-face) may affect the accuracy of self-report. Later studies (Vernon et al, 2008; Beebe et al, 2014), however, found little variation in validation rates according to the mode by which the survey was administered. They thus concluded that mode of data collection results in little potential error.

Test characteristics

Differences in survey format can be a potential contributor to overreporting in validation studies (Vernon, Meissner et al, 2004; Zapka, 2008). These authors call for transparency in *operational definitions* and *measurement construction* to promote comparability across studies. Such transparency will facilitate comparisons of findings. If methods are transparent, even if not

necessarily consistent (e.g., measuring the screening date in terms of month and day vs. within the recommended time interval), comparisons of findings across studies will be facilitated. Standardization in terms of reporting study *results* (i.e., agreement regarding conceptual and operational definitions of behavioral outcomes) should also be a goal (Zapka, 2008).

Item ordering and context need consideration and additional testing as well, and may vary by survey administration mode (Zapka, 2008). A concurrent study by Beebe et al (2008) considered these questions. These authors found that asking about intention to be screened *before* asking about actual recent screening behavior significantly reduced self-reported claims of recent screening. The authors speculated that this likely was due to lowered pressure to give a socially-desirable response. However, self-reports in this study were indeed only *claims* as the authors did not validate them against medical records.

Two of the same authors from this earlier study conducted another recent evaluation (Beebe et al, 2014) that investigated the influence of item ordering and survey administration mode on accuracy of self-report. They included specific descriptions of the different CRC screening tests and compared self-reports with medical records data. This time they concluded that including intention to screen items or survey administration mode (mail vs. telephone) had little impact on CRC screening accuracy.

Medical Records-related Variability

Setting

The uneven quality and feasibility of record access in different health care settings poses a major challenge to CRC validation studies (Zapka, 2008). This author points out that several CRC screening test options (CSPY, DCBE) are performed in specialty practice settings, making it more difficult to ensure complete access to all sources of CRC screening data.

Vernon, Briss et al (2004) suggest that the settings in which studies are conducted may affect validation estimates, particularly the R2R. They assert that this is because studies conducted in clinics often validate only one screening test, usually the most recent one, while those conducted within managed care or health maintenance organizations (HMOs) often attempt to validate more than one serial screening test, e.g., FOBT followed by CSPY. Thus, in studies conducted in this kind of setting, participants may be restricted to those who have had a longer period of membership in the HMO (e.g., at least five years). The highest rates of overreporting, these authors conclude from their review, are found in county health department and public clinic populations (and therefore the ethnic populations that may frequent them) and tumor registries (Vernon, Briss et al, 2004).

Record source accuracy

There are limits to all three data sources – self-reported as well as Medicare claims and medical records. Ferrante et al (2008) emphasize that most studies measuring validity of patient self-reported data assume that the medical record is the "gold standard" and fail to consider the potential errors in documentation and abstraction that regularly occur.

Vernon, Briss et al (2004) point out that medical records kept by managed care programs may be more complete than records kept by public health departments or clinics. As previously discussed, patients in managed care programs usually receive most of their medical care from these same programs, while individuals who receive their medical care from county health departments or public health clinics often do not use the services at these settings consistently for routine care. Thus biases in R2R may be due to incomplete medical records rather than to overreporting by patients.

The quality of documentation of true screening history also influences estimates of specificity. Zapka (2008) suggests that an emphasis on improvement in information technologies may increase the quality of health system and practice data, therefore leading to more accurate screening histories.

FOBT not well documented

Historically, a major problem identified by Sudman et al (1994) involves the availability of test results from FOBTs during medical records abstraction. According to their review, large numbers of test results in paper form are never filed properly in patients' charts. They speculate that this may be due to the lack of the patient identification (ID) number being routinely recorded on the test cards that patients mail back to their offices, thus making it difficult to properly file the information containing the results once they are known. Lipkus et al (2003) confirmed that this kind of CRC screening test was not well documented in the medical records.

Screening vs. diagnostic CRC tests

Identifying those patients who had CRC tests for screening, rather than diagnostic, reasons can be difficult to determine from medical records. CSPY, for example, is both a screening *and* diagnostic CRC test, and, though initially the test may have been ordered for screening purposes, once a polyp is found, the test is reclassified as diagnostic for billing purposes. Vernon et al (2008) conducted an ancillary analysis which excluded patients with gastrointestinal conditions for which the CRC test was ordered, i.e., those who underwent the test for diagnostic purposes. Results of this separate analysis, however, suggest that validity estimates were only minimally affected.

Analysis-related Variability

Liberal vs. strict time interval

Results from the study by Partin et al (2008) suggest that using a more liberal time interval definition, e.g., FOBT within 15 months or CSPY within 11 years of the survey completion date, minimizes the estimated extent of overreporting for overall CRC adherence, thus resulting in improved validity estimates. They conclude that the effect of using a liberal time interval estimate may be preferable when assessing concordance rates as well as more realistic in terms of taking into account waiting times before a patient can be scheduled for a screening. (Such an approach, of course, would result in the ascertainment of screenings that are not truly up-to-date in accordance with USPSTF recommendations.)

Completeness of records

To increase response rates, Shokar et al (2011) attempted to minimize the effect of patients receiving treatment from several offices over time, first, by maximizing the ascertainment of completed tests through the use of multiple sources of chart data and, second, by restricting inclusion criteria to those patients who were in the health system for a sufficient length of time, i.e., the shorter period of "since at least age 50" or "for the past 10 years", to meet the guidelines for all of the CRC tests. (This approach, by definition, results in a reduced number of patients available for consideration.)

Indeterminate data classification

How uncertain and missing data are handled is a potential source of variability in validity studies and hampers comparisons across studies. Partin et al (2008) conclude that when uncertain and missing responses are treated as nonadherent, the extent of overreporting is minimized, albeit not at a statistically significant level. Vernon et al (2008) found that sensitivity and specificity

estimates were minimally affected when respondents with missing values were excluded from analyses, but that the effect of these exclusions was to consistently increase sensitivity while decreasing specificity. These authors note that this effect on sensitivity and specificity estimates could be problematic in studies that have a high proportion of uncertain and missing responses.

The importance of classification of missing or uncertain responses is underscored by results of a study by Griffin et al (2009), based on data from the parent study of veterans by Partin et al (2008) (see Table A2). Though they found that men overreport CSPY more than women, these gender differences were reduced when missing self-reported data were coded as indicative of nonadherence to screening standards rather than as missing data to be excluded. When examining the same individuals' medical records data, they found no gender differences for CSPY. These authors conclude that some of the gender differences found in prior studies using self-reported data only might reflect differences in reporting rather than in screening behavior between males and females and that, moreover, the magnitude of those differences may be reflective of the different ways missing self-report data are coded.

Patient Characteristic-related Variability

Measures of preventive behaviors are particularly problematic across cultures. CRC screening is a particularly complex behavior with multiple influences including personal characteristics, health insurance coverage and patient-physician communication (Beydoun and Beydoun, 2008). The growing cultural diversity of populations increases the complexity of the cognitive tasks involved in self-report (Bastani et al, 2008). Differences among groups can affect the development of survey items at any stage in the process (Shokar et al, 2011). Vernon, Briss et al (2004) point out that cultural differences may affect recall strategies or response editing behaviors and thereby contribute to overreporting.

Such differences in the way time is viewed (e.g., dates and schedules) and in how participants understand which cancer screening test is being discussed may lead to a reduction in the accuracy of self-reported data.

The meta-analysis of Rauscher et al (2008) found lower accuracy of self-reported health behaviors, including cancer screening, among those of minority race as compared to whites, and they suggest that it is possible that similar disparities exist in the accuracy of self-reported CRC screening adherence between men and women. Griffin et al (2009), who examined gender differences in self-reported CRC screening adherence, hypothesized that it is possible that men may confuse prostate screening and CRC screening, thereby leading to inflated differences in reported adherence for men.

One implication of variation in population subgroups for validity estimates is that this variation can mask disparities in prevalence estimates (Bastani et al, 2008). Shokar et al (2011) state that it is important to test the accuracy of validity measures in different groups because such self-report measures are often used to evaluate the effect of CRC screening studies and programs. Results from their study reveal that observed racial/ethnic group differences in test use are real, especially for FOBT, FSIG and DCBE.

Shokar et al (2011) suggest that national prevalence estimates may be overstating screening rates in all groups. Findings of the meta-analysis by Rauscher et al (2008) emphasize the policy implications: when prevalence estimates from national surveys were adjusted using random effects summary estimates, population prevalence estimates decreased, while race/ethnicity disparities increased.

Variability Related to Interactions of Patient Characteristics

Vernon, Briss et al (2004) suggest that, perhaps more importantly, explanations for the overreporting of CRC screening adherence may be

influenced by *interactions* among the various factors discussed above. An apparent gender effect in recall, for example, may be due to other sociodemographic variables (e.g., age, race or educational level) that are different for one gender group vs. the other. These effects can be controlled in the statistical analyses of studies. Use of a different study design, such as one that matches participants on these variables in order to test for recall differences, could also be illustrative (Griffin et al, 2009).

Few studies of adherence to CRC screening have taken into consideration modification of the effect of health factors by sociodemographic factors. Small sample sizes often preclude such control for population group differences in race, education, income, health status or other sociodemographic variables that may have influenced study findings. One study that did examine the possible interaction effects was by Shokar et al (2011), who found that analyses stratified by educational level did not show any differences in sensitivity measures in participants of White and African-American race and Hispanic ethnicity.

Secondary Validity Analysis Studies

Table A2 ("Secondary Validity Analysis Studies") summarizes studies investigating validity estimates (n = 9) stratified by patient sociodemographic characteristics or those identifying correlates of accurate self-reporting through regression analysis. Sociodemographic variables included those commonly reported (e.g., gender, race/ethnicity, age, and education) and identified in a majority of the nine studies as well as others investigated by only more recent studies (e.g., provider recommendation of specific CRC screening, recent visit to provider). Of those studies utilizing regression analysis (four of nine studies), the reported outcome measure ("accuracy") varied from estimates of concordance (Lipkus et al, 2003) to R2R (Hall et al, 2004) to false positive response (Fiscella et al, 2006) to positive predictive value (Reiter et al, 2013).

Results of Analysis of the Secondary Validity Literature

Gender, age, education and family history of CRC were not significantly associated with accurate self-reporting in the majority of the studies that included these data. When these variables were significant, accuracy was generally greater for females than for males (Baier et al, 2000; Hall et al, 2004; Griffin et al, 2009), younger than for older (Lipkus et al, 2003; Hall et al, 2004; Partin et al, 2008), more educated than for less educated (Partin et al, 2008), and those with a family history of CRC than for those without (Partin et al, 2008). Income and insurance coverage were not significantly associated with accuracy in any study.

For race/ethnicity, provider recommendation and recent visit to provider, published results were split between non-significance and significance across varying tests and validity estimates. When significant, whites were generally more accurate than non-whites (Baier et al, 2000; Hall et al, 2004; Fiscella et al, 2006; Shokar et al, 2011), and those with visits to their providers within the past two years were generally more accurate compared to those who hadn't seen their provider as recently (Reiter et al, 2013). One study showing significant results related to provider recommendation was inconclusive in the direction of the reported estimates, with sensitivity estimates excellent for those who said their provider had recommended CRC screening as compared to those without such a recommendation, while specificity estimates were excellent for those who reported *not* having a provider recommendation for CRC testing compared to those who had (White et al, 2013).

Research Questions

An investigation of adherence to CRC screening, taking into account modification of the effect of health factors by sociodemographic factors, is thus needed in multiple population subgroups. This thesis project sought to address

this gap in the literature by examining the relationship between sociodemographic and healthcare utilization variables, including their potential modifying effects, and accuracy of CRC screening self-report in patients residing in diverse geographic locales, primarily within the state of Iowa. It addresses three main research questions.

Primary Research Question

"Is patient self-report of up-to-date CRC screening status, using the medical record as the reference standard, accurate?"

This will be addressed in the primary data analysis.

Secondary Research Questions

"Are sociodemographic and healthcare utilization factors associated with accuracy, when defined as concordance or agreement between self-report and medical record documentation, for up-to-date (UTD) CRC screening by CSPY?"

"Does adjustment for these factors or covariates modify such an association?"

These questions will be addressed in the secondary data analysis.

METHODS

Sample

Study participants were patients at 16 family medicine offices that are part of the Iowa Research Network (IRENE), a practice-based research network (PBRN). The 16 offices were randomly selected from 27 family medicine offices which initially had provided letters of support for the grant application of the parent study, a randomized controlled trial of educational interventions to improve CRC screening rates in IRENE offices (Levy et al, 2012). The parent study, with accompanying statistical analyses such as those involved in this project, was approved by the Institutional Review Board of the University of Iowa. Data were collected between 2010 and 2011. Participating offices were located in communities of varying population size in rural counties with a median per capita income level below the state average. Fourteen were located in the state of Iowa; two were located just over the border in neighboring South Dakota.

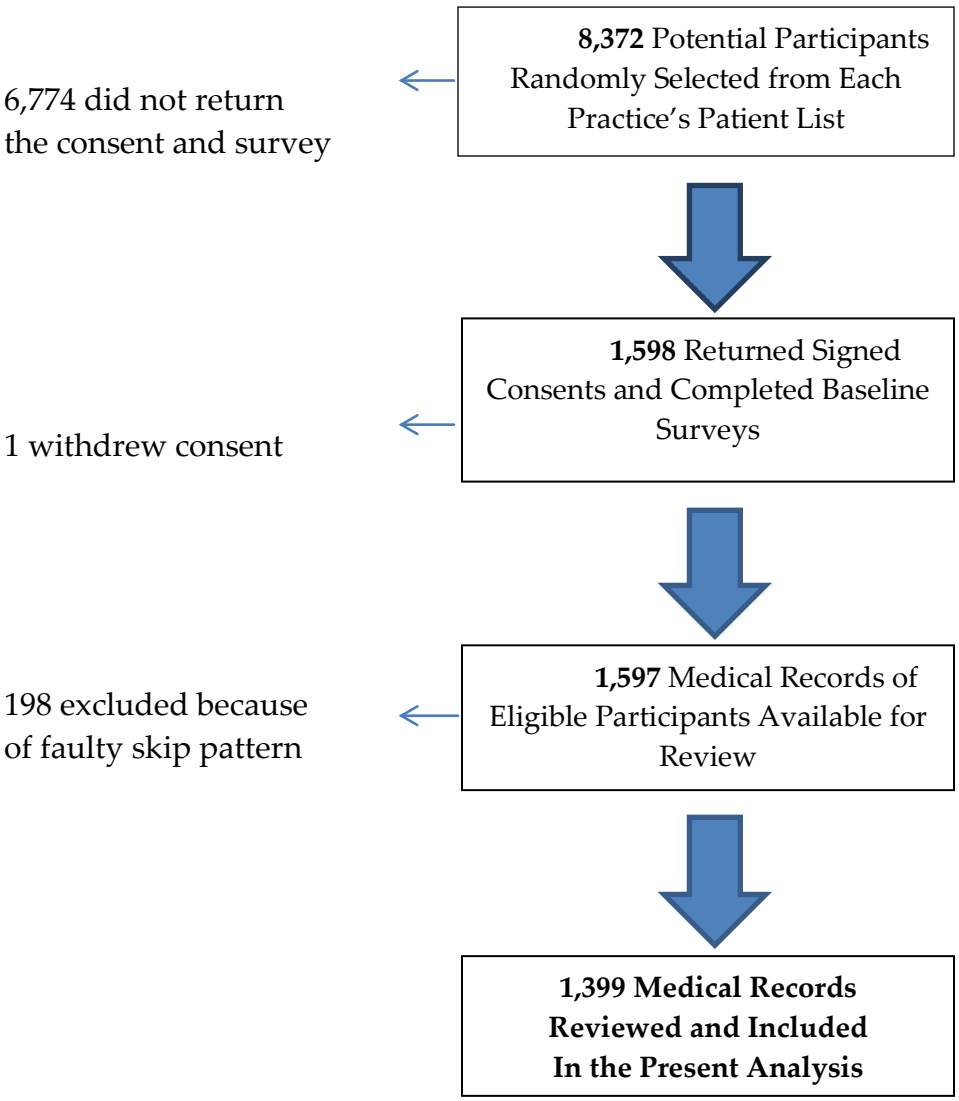
Design

Full details of the parent study design are described elsewhere (Daly et al, 2012). Of significance to this validity study, each office provided a list of patients, originally specified to be aged 52 to 79 years. From these lists, individuals residing in nursing homes were excluded from consideration because of concerns about overall health status and potential inability to provide informed consent. A sample of 530 individuals (265 males and 265 females) was randomly selected from each of the 14 larger offices' lists and invited to participate in the parent study; two smaller offices did not have a sufficient patient population to meet the desired sample size specifications and thus provided a list of fewer individuals, all of whom were invited to participate in the parent study.

Recruitment procedures were adapted from the Dillman approach (Dillman, 2000) to optimize response to self-administered mail surveys. A total of 8,372 individuals were invited, via an introductory letter sent by mail, to participate in the CRC screening intervention study. This was followed by a second mailing two to three weeks later which included an informed consent document (Appendix B), a baseline survey (Appendix C) and a $2 bill as cash incentive. If no response was received after two to three weeks of the second mailing, a reminder letter was sent, followed by a phone call (with up to four attempts at various times of the day) to non-responders two to six weeks after the reminder letter was mailed. If there still was no response from an invited individual, a final mailing of the consent documents and baseline survey was sent.

Of 8,372 potential participants, 1,598 (19%) interested individuals returned completed baseline surveys as well as signed consents in which they agreed to participate in the study and provided permission for study personnel to access medical records at the individual's IRENE office (Figure 1). One participant elected to withdraw consent after initially agreeing to participate; this left 1,597 participants who completed baseline surveys and also had medical records at their IRENE office available for analysis. A faulty skip pattern in the initial version of the survey instrument (see Appendix C, Question 9: respondents who answered "no" to this question were instructed to "go to question 20" rather than on to question 10) prevented some participants from responding to questions about completion of certain CRC tests, resulting in 198 (12%) participants' surveys being excluded for later analysis. In the end, 1,399 participants, or 17% of those initially invited, were included in this analysis.

Figure 1. Subject Flow Diagram.



Data Collection

Interested individuals self-completed an investigator-developed baseline survey (see Appendix C) which recorded sociodemographic characteristics, personal and family history of CRC, and personal history of and attitudes toward, including barriers to and readiness for, CRC screening. In addition, the baseline survey queried participants regarding their past CRC screening with specific questions for each of the major tests. For each test, they were first asked if their doctor had recommended the test. If they responded affirmatively, they were asked if they completed the test, and, if so, if this was within the recommended time interval based on USPSTF guidelines in effect at the time for average risk individuals between the ages of 50 and 75 years: either a fecal occult blood test (FOBT) or fecal immunochemical test (FIT) annually, flexible sigmoidoscopy (FSIG) or double contrast barium enema (DCBE) every five years, colonoscopy (CSPY) every 10 years, or a combination of annual FOBT/FIT and FSIG every five years (U.S. Preventive Services Task Force, 2008).

By design, these intervals were set up to correspond to those recommended in national CRC screening guidelines, such as the USPSTF, ACS (Levin, Lieberman, McFarland, Smith et al, 2008) and AGA (Levin, Lieberman, McFarland, Andrews et al, 2008), which remain the most recently published. They did not require individuals to specify exact month or year of testing. If participants responded negatively or were unsure if their doctor had recommended the test, they were instructed to skip to the next set of questions, asking about a different CRC screening test.

Approximately 15 months after the initial introductory letter was mailed, post-enrollment chart reviews were conducted for all consenting patients by trained research team members who traveled to each of the 16 IRENE offices. Records were reviewed to identify completion dates for each CRC screening test,

along with dates of the patient's first and most recent visit to the office. Eight of the 16 offices utilized electronic medical records (EMRs), and abstraction proceeded following brief instruction with regard to their navigation; for the remainder of the offices, paper charts were in use at the time of the abstraction.

All relevant outpatient notes and laboratory, imaging and ancillary reports were reviewed by a research team member using a standardized abstraction form. The research team member was blinded to information in the participant's self-reported survey. Approximately 10% of these records were reviewed by a second research team member for accuracy of abstracting. An expert physician from the study team was consulted at the time of the review to resolve any disagreements.

Because many consented individuals in the study had CRC screening tests performed by healthcare providers outside of their IRENE office, specifically, by specialist gastroenterologists who performed CSPY over the relevant time period, CRC screening was requested from these other providers. For a subset (n = 35) of patients, a note or letter in the medical record suggested that these patients had had a CRC screening test within the specified time period. This information often included the physician and/or hospital where the procedure had been performed but not confirmatory documentation, e.g., the CSPY report. For this subset, the participant's written authorization was secured to access these outside records. Information requested from outside facilities included the screening test (e.g., CSPY) and associated pathology reports, if applicable. Out of the 35 requests, tests and/or reports were returned for 34 patients.

Calculated and Derived Variables

In addition to the sociodemographic variables gathered from the self-reported baseline survey, two involving healthcare utilization were calculated from data obtained from the medical record; a third, the geographic variable,

involved classification according to the sample patient addresses provided by the 16 IRENE offices.

Healthcare Utilization Variables

Recency of Last Visit. This was defined as the difference between the date of medical record review and date of patient's last visit to the office.

Duration of Patient Status. This was defined as the difference between the date of medical record review and date of patient's first visit to the office.

Geographic Variable

RUCA Category. This rurality classification of patient residence is based on the Rural-Urban Commuting Area (RUCA) classification scheme developed by the Office of Rural Health Policy (Hart et al, 2005). The scheme classifies areas based on census tract, or in the case of the present study, zip code approximation, taking into account population as well as work-commuting information. Patient addresses were assigned a RUCA code by zip code. RUCA codes were grouped into four categories of "Metropolitan" (population of 50,000+), "Large Rural" (population 10,000-49,999), "Small Town" (2,500-9,999) and "Rural Area" (< 2,500), according to the commonly-used RUCA Categorization A (WWAMI Rural Health Research Center, retrieved April 27, 2014).

Basis for Validity Estimates (Outcome Measures)

Participants were considered adherent if there was documentation of an FOBT/FIT within 12 months, an FSIG or DCBE within 5 years or a CSPY within 10 years prior to the baseline survey completion date. If records indicated multiple up-to-date tests of the same type, the most recent test was used. Overall CRC screening adherence was defined as adherence to any one of the tests ("any") within the specified time interval, as defined by the USPSTF guidelines (U.S. Preventive Services Task Force, 2008).

Two items from the survey instrument were used to assess self-reported CRC screening adherence status for each test: (1) a yes/no question asking if the patient had ever had the particular test in question (or, with wording specific to the FOBT/FIT test, "returned" the test) and (2) a question asking how long ago the patient had had the test, with answers framed according to the test-specific adherence standards (e.g., "within the past 10 years" or "more than 10 years ago" for CSPY). The first question, asking if the patient had ever had a particular test, was preceded by an initial question asking if the patient's doctor had ever recommended that test, as it was assumed that the patient would have to have had a physician recommendation in order to complete the test. If the patient responded "no" or "unsure" to this question, they were instructed to skip the subsequent two items, described above, upon which adherence was based. If, however, they chose to ignore this skip pattern and proceeded to answer the two items described above, their responses were included.

If patients responded "yes" to having had a CRC screening test and answered that the test was completed within the recommended time interval for that particular procedure (e.g., "within the past year" for FOBT/FIT), they were coded as adherent based on self-report. If they did not answer the initial yes/no question or answered "unsure", they were assigned indeterminate status, regardless of whether the recency question was answered. If patients answered the first question affirmatively but did not answer the recency question, their responses were considered indeterminate. Indeterminate responses were considered in two ways, either as missing data to be excluded or as evidence of nonadherence, and the results of the two different classification schemes were compared in additional analyses.

Primary Analyses

The four validity measures, sensitivity, specificity, concordance and report-to-records ratio (R2R), were used to summarize the validity of self-reported CRC screening adherence for "any" test, i.e., up-to-date by any one of the four CRC screening tests assessed (FOBT/FIT, CSPY, FSIG, DCBE) and for each individual test. Two-sided 95% confidence intervals were calculated around these estimates using standard methods for proportions (for sensitivity, specificity and concordance estimates) and comparison of two proportions (for R2R), similar to the population relative risk (Pagano and Gauvreau, 2000).

Ancillary Analyses

Liberal vs. strict time interval

Similar to the approach of Partin et al (2008) regarding time interval, an additional ancillary analysis used a liberal vs. strict time interval when assessing validity estimates comparing self-report and medical records data. The strict time interval followed USPSTF recommendations, while the liberal definition was 15 months for FOBT/FIT, 5.5 years for FSIG or DCBE and 11 years for CSPY.

Indeterminate data classification

Based on the findings of Partin et al (2008), a separate analysis compared results of the full sample, defining indeterminate cases as *nonadherent*, to a subset which *excluded* them.

Completeness of records

Following the procedures of Shokar et al (2011), to measure the effect of completeness of medical records in this study, validity estimates of participants who were patients in their IRENE office for at least 10 years were compared with those from the entire sample, some of whom may have received outside care and for whom records could not be obtained.

Secondary Analyses

To assess associations that might explain variation in validity estimates between self-report and medical records data, univariate and multivariable analyses using logistic regression methods were conducted. Predictor (independent) variables included the following, with categories from the baseline survey listed in parentheses: age; gender; education (8th grade or less, some high school, high school graduate or General Education Diploma (GED), some college or technical school, college graduate, graduate degree); annual household income (less than $20,000, $20,000-$39,999, $40,000-59,999, $60,000-$79,999, $80,000-$99,999, greater than $100,000); insurance status (individual plan/self-pay, group plan through employer, government plan, e.g., military-service connected or through a Veterans Affairs (VA) facility, Medicaid, Medicare, Medicare plus self-pay supplement, no insurance); immediate family, i.e., parent or sibling, history of CRC; and IRENE office attended. Additional predictor variables included the three calculated variables, recency of last visit to office, duration of patient status in IRENE office and RUCA category based on patient zip code. Wald chi-square statistics and associated p-values, along with odds ratios and Wald 95% confidence intervals, were used for inference.

Concordance, or agreement between self-report on baseline survey and medical records data, for CSPY was the outcome variable for logistic regression analyses. This CRC screening test was chosen as the one assessed in the regression analyses because it was the most common test documented in the medical record (MR).

Univariate analysis of each predictor variable was initially conducted, beginning with construction of contingency tables, fitting of univariate logistic models, estimation of odds ratios and testing for significant associations. Determination of reference groups for discrete variables included consideration

of useful comparisons, e.g., predicted highest or lowest category, in a well-defined group, with a reasonably large sample size compared with other groups.

To select variables for the multivariable analysis, those with Wald Chi-square p-values of < 0.25 in the univariate analysis, along with others considered clinically relevant, were included, and models were fitted by removing variables, manually, from the full model in a backward sequential fashion. The importance of each variable was verified through examination of the Wald statistic, comparing each estimated coefficient to the coefficient from the univariate model. Any variable which did not lead to a 15 percent change in the parameter estimates was eliminated (Hosmer and Lemeshow, 2000). This process of "delete-refit-verify" was continued until all important variables were included in the model, and all excluded variables were considered either clinically or statistically unimportant, producing a preliminary main effects model.

Variables in this preliminary main effects model were examined more closely to consider the assumption of linearity in the logit and the categorization used for discrete variables. Those variables found to be non-linear were transformed into a categorical form, and categories of some discrete variables with similar estimates were collapsed. Thus, in this main effects model, the best form for those variables identified as important in the multivariable model was determined.

Finally, possible interaction effects among the variables in the multivariable model were considered. The contribution of each interaction effect to the multivariable model was assessed, one at a time, and any that were evaluated as either clinically or statistically significant were included in the final fitted multivariable model. Model fit statistics were evaluated to determine the adequacy of the final model.

A priori Alternative Hypotheses

Based on the literature review (Appendices A and B), the following *a priori* hypotheses were specified:

Alternative Hypotheses – Primary Analyses

1) Sensitivity, specificity and concordance estimates will be acceptable (≥ 0.7, based on the criteria of Tisnado et al, 2006), for "any" test and for each individual test (FOBT/FIT, CSPY, FSIG and DCBE).

2) Overreporting (R2R > 1.00) will be found for "any" test and for each individual test.

3) CSPY estimates will be more accurate than those for other CRC screening tests.

4) Use of a more liberal time interval within which screening occurred will result in more favorable estimates.

5) Excluding patients with indeterminate data from the analyses, rather than treating such data as evidence of nonadherence, will result in less favorable validity estimates.

Alternative Hypotheses – Secondary Analyses

1) Patients with fewer years of education will be less accurate in reporting screening status than those with more years of education.

2) Patients who have an immediate family history of CRC will be more accurate in reporting screening status than those who have no family history of CRC.

3) Patients who reside in more rural settings, as determined by RUCA category classification, will be less accurate in reporting screening status than those who reside in more urban settings.

RESULTS

Description of the Sample

Sample characteristics for the 1,399 patients whose medical records were reviewed are presented in Table 1. The mean age was 62.6 years, with a standard deviation (SD) of 7.1. This is similar to the mean age of 63.3 years for non-responders and those otherwise not included in the final analysis because of withdrawal of consent and the faulty survey skip pattern. Approximately 53% of the final sample was female as compared to 50% for non-responders and those otherwise not included. Geographically, responders were similar to non-responders. About 30% of patients were classified as residing in a "rural area" (comparable to non-responders), almost one-quarter were from a "small town" (also comparable to non-responders), just over one-quarter were from a "large rural" area (as compared to just under one-quarter for non-responders), and approximately one in five was from a "metropolitan" area (as compared to slightly more than one in five for non-responders).

Approximately 99% were non-Hispanic and white. Seventy-nine percent were married, and about 65% had an education beyond high school. Almost 64% reported an annual household income of $40,000 or more. Almost all had some form of insurance, nearly one-half through their employer and more than one-third from Medicare, either with or without a private supplement. More than one in four reported a first-degree relative with CRC. The mean time since patients' most recent physician visit was 0.9 years – not quite 11 months – while the average length of time that participants were patients in their IRENE offices was just over 17 years. Nearly four out of five reported that their physician had recommended they undergo CRC screening. Response rate per clinic ranged from 8 to 46% with a mean of 18%.

Table 1. Sample Characteristics for Patients Included in the Analysis, Iowa Research Network (IRENE), 2010-2011

Sociodemographic Variables (N = 1,399)	mean	SD
Age, years	62.6	(7.1)
	n	(%)
Gender		
Female	744	(53.2)
Male	655	(46.8)
Race		
White	1372	(98.1)
non-White	16	(1.1)
Missing	11	(0.8)
Ethnicity		
Hispanic	11	(0.8)
non-Hispanic	1347	(96.3)
Missing	41	(2.9)
Marital Status		
Married/marriage-like relationship	1098	(78.5)
Single	168	(12.0)
Widowed	121	(8.6)
Missing	12	(0.9)
Education		
8th grade or less	14	(1.0)
Some high school	53	(3.8)
High school diploma/GED	414	(29.6)
Some college/technical school	482	(34.5)
Bachelor's degree	245	(17.5)
Post-graduate degree	176	(12.6)
Missing	15	(1.0)

Table 1. Continued

Annual Household Income		
< $20,000	164	(11.7)
$20,000-$39,999	307	(21.9)
$40,000-$59,999	330	(23.6)
$60,000-$79,999	250	(17.9)
$80,000-$99,999	124	(8.9)
$100,000+	118	(8.4)
Missing	106	(7.6)
Insurance		
Individual (self-pay)	158	(11.3)
Group (through employer)	618	(44.2)
Military service-connected	32	(2.3)
Medicaid	17	(1.2)
Medicare	161	(11.5)
Medicare with self-pay supplement	352	(25.2)
None	43	(3.1)
Missing	18	(1.3)
Family History (first-degree relative) of Colorectal Cancer (CRC)		
Yes	384	(27.4)
No or unsure	1015	(72.6)
Rural-Urban Commuting Area (RUCA) category		
Metropolitan	275	(19.7)
large rural	373	(26.7)
small town	324	(23.2)
rural area	427	(30.5)
Physician recommended CRC screening		
Yes	1066	(76.2)
No or unsure	289	(20.7)
Missing	44	(3.1)

Table 1. Continued

IRENE Office		
Alcester	73	(5.2)
Algona	93	(6.6)
Clinton	109	(7.8)
Corning	150	(10.7)
Dubuque	47	(3.4)
Elk Point	76	(5.4)
Grinnell	74	(5.3)
Guttenberg	73	(5.2)
Iowa Falls	58	(4.1)
Le Mars	61	(4.4)
Manchester	42	(3.0)
Muscatine	84	(6.0)
Sioux Center	92	(6.6)
Sioux City	151	(10.8)
Spencer	82	(5.9)
West Burlington	134	(9.6)
Healthcare Utilization Variables	mean	SD
Most recent patient visit, years (n = 1,335)	0.9	(1.55)
Patient duration in office, years (n = 1,295)	17.1	(11.45)

Table 2 provides the raw data from which all primary validity estimates were derived. Sixty percent (837/1399) of patients reported they were up-to-date with CRC screening by any test; this compares to 48% (675/1399) that had such documentation in the MR. About one-quarter of those who said they were up-to-date by any test did not have evidence of this in the MR. Of those with documentation of a test in the MR, 95% were screened by CSPY. Few patients had FOBT/FIT, FSIG or DCBE test results in the MR.

Results of the Primary Analyses

Table 3 presents validity estimates for "any" test and for FOBT/FIT, CSPY, FSIG and DCBE, utilizing a strict time interval and treating indeterminate data as nonadherent. Using the criteria proposed by Tisnado et al (2006) to evaluate the results, sensitivity was excellent for "any" test and CSPY, good for DCBE and poor for FOBT and FSIG. Specificity was excellent for FOBT/FIT, FSIG and DCBE and fair for "any" and CSPY. Concordance was excellent for FOBT, FSIG and DCBE and good for "any" and CSPY. R2R ranged from 1.21 (CSPY) to 7.10 (FSIG).

Results of the Ancillary Analyses

Table 4 provides estimates using strict vs. liberal time intervals and when classifying indeterminate data as nonadherent (first two sections) vs. excluded (final two sections). Comparison of the first and second sections of Table 4 with the third and fourth sections shows that estimates are very similar, indicating that using a strict or more liberal time interval made little difference. Sensitivity remained excellent for "any" test and CSPY and poor for FOBT/FIT and FSIG, regardless of time interval or treatment of indeterminate data. For DCBE, sensitivity improved from good to perfect (1.00) when indeterminate data were excluded, regardless of the time interval. Specificity estimates were identical when comparing strict and liberal time intervals and treating indeterminate data

Table 2. Cross-classification of Adherence to Colorectal Cancer Screening Based on Self-report and Medical Record Review, Iowa Research Network (IRENE), 2010-2011

			Medical Record: Adherent	Medical Record: Nonadherent	
Self-report					
"Any"[1]	Adherent[2]	TP	640	197	FP
	Nonadherent	FN	3	22	TN
	Indeterminate		32	505	
FOBT	Adherent	TP	12	78	FP
	Nonadherent	FN	7	521	TN
	Indeterminate		1	780	
CSPY	Adherent	TP	624	177	FP
	Nonadherent	FN	18	309	TN
	Indeterminate		22	249	
FSIG	Adherent	TP	1	70	FP
	Nonadherent	FN	4	436	TN
	Indeterminate		5	883	
DCBE	Adherent	TP	5	36	FP
	Nonadherent	FN	0	354	TN
	Indeterminate		1	1003	

[1] "Any" is representative of overall screening adherence.

[2] Patients were classified as adherent if they reported completing either a fecal occult blood test (FOBT) within 12 months of receipt of survey, a flexible sigmoidoscopy (FSIG) or barium enema (DCBE) within 5 years of receipt of survey, or a colonoscopy (CSPY) within 10 years of receipt of survey. For the primary analyses, indeterminate responses ("unsure" or missing) were combined with nonadherent, as indicated by the blue brackets.

Table 3. Comparison of Adherence to Colorectal Cancer Screening Based on Self-report with Adherence Based on Medical Record, Iowa Research Network (IRENE) Patients (N = 1,399), 2010-2011

	Self-report		Medical Record	
	n	(%)	*n*	(%)
Screening Test[1]				
"Any"[2]	837	(59.8)	675	(48.2)
FOBT	90	(6.4)	20	(1.4)
CSPY	801	(57.3)	664	(47.5)
FSIG	71	(5.1)	10	(0.7)
DCBE	41	(2.9)	6	(0.0)

[1] Patients were classified as adherent if they reported completing either a fecal occult blood test (FOBT) within 12 months of receipt of survey, a flexible sigmoidoscopy (FSIG) or barium enema (DCBE) within 5 years of receipt of survey or a colonoscopy (CSPY) within 10 years of receipt of survey.

[2] "Any" is representative of overall screening adherence.

Table 4. Comparison of Results of Primary and Ancillary Analyses for "Any" Test and Test-specific Validity Estimates and 95% Confidence Intervals, Iowa Research Network (IRENE) Patients, 2010-2011

	Sensitivity		Specificity		Concordance		R2R	
Strict[1], indeterminate as nonadherent								
"Any"[2] test	0.95	(0.93-0.96)	0.73	(0.69-0.76)	0.83	(0.81-0.85)	1.24	(1.19-1.29)
Fecal Occult Blood Test	0.60	(0.38-0.78)	0.94	(0.93-0.95)	0.94	(0.92-0.95)	4.50	(2.38-8.51)
Colonoscopy	0.94	(0.92-0.96)	0.76	(0.73-0.79)	0.84	(0.82-0.86)	1.21	(1.16-1.26)
Flexible Sigmoidoscopy	0.10	(0.02-0.41)	0.95	(0.94-0.96)	0.94	(0.93-0.95)	7.10	(0.48-104.76)
Digital Contrast Barium Enema	0.83	(0.42-0.96)	0.97	(0.96-0.98)	0.97	(0.96-0.98)	6.83	(2.79-16.74)
Liberal[2], indeterminate as nonadherent								
"Any" test	0.95	(0.93-0.96)	0.73	(0.70-0.76)	0.83	(0.81-0.85)	1.23	(1.18-1.29)
Fecal Occult Blood Test	0.54	(0.35-0.71)	0.94	(0.93-0.96)	0.94	(0.92-0.95)	3.46	(1.90-6.30)
Colonoscopy	0.94	(0.91-0.95)	0.76	(0.73-0.79)	0.85	(0.82-0.86)	1.20	(1.15-1.25)
Flexible Sigmoidoscopy	0.10	(0.02-0.41)	0.95	(0.94-0.96)	0.94	(0.93-0.95)	7.10	(0.48-104.76)
Digital Contrast Barium Enema	0.83	(0.42-0.96)	0.97	(0.96-0.98)	0.97	(0.96-0.98)	6.83	(2.79-16.74)
Strict, indeterminate excluded								
"Any" test	1.00	(0.99 - 1.00)	0.10	(0.07 - 0.15)	0.77	(0.74 - 0.79)	1.30	(1.25-1.35)
Fecal Occult Blood Test	0.63	(0.41 - 0.81)	0.87	(0.84 - 0.89)	0.86	(0.83 - 0.89)	4.74	(2.53-8.88)
Colonoscopy	0.97	(0.96 - 0.98)	0.64	(0.59 - 0.68)	0.83	(0.80 - 0.85)	1.25	(1.20-1.30)
Flexible Sigmoidoscopy	0.20	(0.04 - 0.64)	0.86	(0.83 - 0.89)	0.86	(0.82 - 0.88)	14.20	(1.03-194.90)
Digital Contrast Barium Enema	1.00	(0.54 - 1.00)	0.91	(0.87 - 0.93)	0.91	(0.88 - 0.93)	8.20	(3.61-18.64)

Table 4. Continued

Liberal, indeterminate excluded								
"Any" test	1.00	(0.99-1.00)	0.10	(0.07-0.15)	0.77	(0.74-0.80)	1.30	(1.25-1.35)
Fecal Occult Blood Test	0.56	(0.37-0.73)	0.87	(0.84-0.89)	0.86	(0.83-0.88)	3.60	(1.99-6.52)
Colonoscopy	0.97	(0.96-0.98)	0.64	(0.60-0.68)	0.83	(0.81-0.85)	1.24	(1.19-1.29)
Flexible Sigmoidoscopy	0.20	(0.04-0.64)	0.86	(0.83-0.89)	0.86	(0.82-0.88)	14.20	(1.03-194.90)
Digital Contrast Barium Enema	1.00	(0.54-1.00)	0.91	(0.88-0.93)	0.91	(0.88-0.93)	8.20	(3.61-18.64)

[1] Strict = Patients were classified as adherent if they reported completing either a fecal occult blood test (FOBT) within 12 months of receipt of survey, a flexible sigmoidoscopy (FSIG) or double contrast barium enema (DCBE) within 5 years, or a colonoscopy within 10 years; [2] Liberal = Patients were classified as adherent if they reported completing a FOBT within 15 months of receipt of survey, a FSIG or DCBE within 5.5 years, or a CSPY within 11 years.

[2] "Any" is representative of overall screening adherence.

as nonadherent (first and second sections). However, for strict or liberal time frame, when indeterminate data were excluded (third and fourth sections), specificity fell from fair to poor for "any" test and CSPY, and from excellent to good for FOBT and FSIG, while DCBE remained good.

Concordance estimates also remained essentially unchanged when comparing time intervals but fell from good to fair for "any" test and from excellent to good for FOBT/FIT and FSIG when indeterminate data were excluded; DCBE remained excellent and CSPY was essentially unchanged, regardless of how indeterminate data were handled. Time interval did not affect R2R with the exception of FOBT, where use of a more liberal time interval increased the ratio. With regard to consideration of indeterminate data, R2R increased for all four tests when such data were excluded; this was most dramatically seen for FSIG and, to a lesser extent, for DCBE.

Table 5 compares estimates from the analysis of Table 3, calculated for the entire sample using a strict time interval and considering indeterminate data as nonadherent, with estimates from those who were patients in their IRENE for at least 10 years. Estimates were essentially unchanged for the two groups.

Results of the Secondary Analyses

A summary of the univariate logistic regression analysis results of concordance between self-report and MR for CSPY is presented in Table 6. Several factors were significantly ($p < 0.05$) associated with concordance for CSPY, including Education ($p = 0.043$), Insurance ($p < 0.001$), Family History of CRC ($p < 0.001$) and RUCA Category ($p = 0.011$).

With "high school graduate" as the reference group, the odds that those with less than a high school education ("some high school") reported accurately

Table 5. Comparison of "Any" Test and Test-specific Validity Estimates and 95% Confidence Intervals for Total Sample (N = 1,399) with Patients in Office 10+ Years (N = 824).

CRC Screening Test[1]	Sensitivity		Specificity		Concordance		R2R	
"Any"[2] test								
Total sample	0.95	(0.93-0.96)	0.73	(0.69-0.76)	0.83	(0.81-0.85)	1.24	(1.19-1.29)
Patients in office 10+ years	0.95	(0.92-0.96)	0.74	(0.69-0.78)	0.85	(0.82-0.87)	1.19	(1.13-1.25)
FOBT								
Total sample	0.60	(0.38-0.78)	0.94	(0.93-0.95)	0.94	(0.92-0.95)	4.50	(2.38-8.51)
Patients in office 10+ years	0.64	(0.38-0.84)	0.94	(0.93-0.96)	0.94	(0.93-0.95)	3.93	(1.92-8.02)
CSPY								
Total sample	0.94	(0.92-0.96)	0.76	(0.73-0.79)	0.84	(0.82-0.86)	1.21	(1.16-1.26)
Patients in office 10+ years	0.93	(0.90-0.95)	0.78	(0.74-0.82)	0.86	(0.83-0.88)	1.14	(1.09-1.20)
FSIG								
Total sample	0.10	(0.02-0.41)	0.95	(0.94-0.96)	0.94	(0.93-0.95)	7.10	(0.48-104.76)
Patients in office 10+ years	0.17	(0.04-0.58)	0.95	(0.94-0.97)	0.95	(0.93-0.96)	6.50	(0.47-90.65)
DCBE								
Total sample	0.83	(0.42-0.96)	0.97	(0.96-0.98)	0.97	(0.96-0.98)	6.83	(2.79-16.74)
Patients in office 10+ years	0.75	(0.28-0.95)	0.98	(0.96-0.99)	0.98	(0.96-0.98)	5.50	(1.67-18.15)

[1] Patients were classified as adherent if they reported completing a fecal occult blood test (FOBT) within 12 months of receipt of survey, a flexible sigmoidoscopy (FSIG) or barium enema (DCBE) within 5 years of receipt of survey, or a colonoscopy (CSPY) within 10 years of receipt of survey.

[2] "Any" is representative of overall screening adherence. Indeterminate (unsure or missing) responses were classified as nonadherent.

Table 6. Univariate Logistic Regression Analysis of Accuracy, Defined as Concordance Between Self-report of Up-to-date Colonoscopy Screening Status and Documentation in Medical Record, Using Categories from the Baseline Survey, Iowa Research Network (IRENE) patients, 2010-2011.

Variable	OR	Wald 95% CI	p-value
Gender (n = 1,399)			0.2100
Male	1.00	reference	
Female	1.20	(0.90-1.61)	0.2100
Education (n = 1,384)			0.0429
Graduate degree	1.17	(0.71-1.92)	0.5446
College degree	1.20	(0.77-1.87)	0.4275
Some college	1.20	(0.83-1.73)	0.3398
High School graduate	1.00	reference	
< High School graduate	0.49	(0.27-0.88)	0.0175
Income[1] (n = 1,293)			0.1299
$100,000+	0.79	(0.45-1.41)	0.4292
$80,000-99,999	1.51	(0.77-2.96)	0.2280
$60,000-79,999	1.10	(0.68-1.79)	0.6905
$40,000-59,999	1.00	reference	
$20,000-39,999	0.73	(0.47-1.11)	0.1401
< $20,000	0.70	(0.42-1.15)	0.1535
Insurance (n = 1,381)			0.0002
Group	1.00	reference	
Medicaid	0.44	(0.14-1.39)	0.1630
Medicare	0.65	(0.40-1.04)	0.0712
Medicare + Supplement	0.53	(0.37-0.75)	0.0004
Military Service-connected	0.26	(0.12-0.56)	0.0006
Individual	0.89	(0.53-1.49)	0.6523
None	5.71	(0.78-42.02)	0.0874
Family History[2] of CRC (n = 1,397)			0.0007
No	1.00	reference	
Yes	0.54	(0.38-0.77)	0.0007

Table 6. Continued

RUCA Category (n = 1,399)			0.0114
Metropolitan	0.71	(0.43-1.16)	0.1706
Large rural	0.49	(0.32-0.76)	0.0014
Small town	1.00	reference	
Rural area	0.58	(0.37-0.89)	0.0138
IRENE Office (n = 1,399)			0.0511
Corning	1.00	reference	
Alcester	0.52	(0.25-1.08)	0.0779
Algona	1.20	(0.53-2.72)	0.6550
Clinton	0.61	(0.31-1.20)	0.1493
Dubuque	0.99	(0.37-2.65)	0.9858
Elk Point	0.70	(0.33-1.51)	0.3675
Grinnell	0.57	(0.27-1.20)	0.1390
Guttenberg	0.74	(0.34-1.62)	0.4461
Iowa Falls	1.96	(0.64-6.02)	0.2414
Le Mars	1.33	(0.50-3.51)	0.5651
Manchester	1.38	(0.44-4.30)	0.5806
Muscatine	0.50	(0.25-1.00)	0.0504
Sioux Center	1.52	(0.64-3.64)	0.3435
Sioux City	1.01	(0.51-1.99)	0.9825
Spencer	0.60	(0.29-1.24)	0.1666
West Burlington	0.53	(0.28-0.99)	0.0461
Age* (n = 1,399)	0.85	(0.77-0.94)	0.0011
Recency of Last Visit** (n = 1,335)	0.77	(0.71-0.84)	< 0.0001
Duration of Patient Status** (n = 1,295)	1.02	(1.01-1.03)	0.0090

[1] Annual household income; [2] First-degree relative.

OR is for an increase of 5 years (*) or 1 year (**).

Abbreviations: CI = confidence interval; CRC = colorectal cancer; OR = odds ratios; RUCA = Rural-Urban Commuting Area.

were significantly lower (OR = 0.49; p = 0.0175). No other education category differed significantly from "high school graduate", although the apparent trend across categories for Education was for increased education to be associated with increased likelihood of concordant reporting. With those covered by a "group" insurance plan, i.e., through an employer or union, as the reference group, the odds that those patients with military service-connected insurance or Medicare plus supplemental insurance they paid for reported accurately were 26% (OR = 0.26) and 53% (OR = 0.53), respectively. Patients who reported having an "individual" plan, i.e., self-pay, those covered by Medicaid and Medicare (without a supplement) as well as those who reported having no insurance in the past year did not differ significantly from those with group insurance.

With "no family history" of CRC in a first-degree relative as the reference group, the odds that those who reported having an immediate family member with CRC history reported accurately were 54% (OR = 0.54). With "small town" as the reference group for RUCA Category, the odds that those who resided in "large rural" and "rural area" categories reported accurately were 49% (OR = 0.49) and 58% (OR = 0.58), respectively. Those who resided in the "metropolitan" category did not differ significantly from those who resided in a small town.

Gender and Income, while not statistically significantly associated with concordance (p = 0.2100 and p = 0.1299, respectively), were considered clinically relevant, and, with Wald Chi-square p-values of < 0.25, were included in the multivariable model building. IRENE Office, with a Wald Chi-square p-value of 0.0511, was not considered for inclusion in the multivariable model because of multicollinearity and sample size concerns. All continuous variables were significantly (p < 0.05) associated with concordance: Age (OR = 0.85, 95% CI = 0.77-0.94, for an increase of five years in age); Recency of Last Visit (OR = 0.77, 95% CI = 0.71-0.84, for an increase of one year since the last visit); and Duration

of Patient Status (OR = 1.02, 95% CI = 1.01-1.03, for an increase of one year of being seen at that IRENE office).

With the exception of IRENE Office, all variables in the univariate analysis were therefore included in the preliminary multivariable main effects model (not shown). Close examination of the continuous variables for the assumption of linearity in the logit resulted in the transformation of two variables (Recency of Last Visit and Duration of Patient Status) into a categorical form; some categories of three discrete variables (Education, Income and Insurance) were collapsed because of small cell sizes if estimates were in a similar direction. Consideration of possible effect modification resulted in the inclusion of two interaction terms (RUCA category x Age and Recency of Last Visit x Age).

Table 7 presents the final multivariable model. All but two variables (Gender and Income) remained in the model. Variables significantly ($p < 0.05$) associated with concordance for CSPY, now adjusted for all other variables in the model, included Education, Insurance, Family History of CRC and Duration of Patient Status.

With "high school graduate" as the reference group, the odds that those with less than a high school education ("some high school") reported accurately were 40% (aOR = 0.40). With "group" insurance as the reference group, the odds that those with Medicare, with or without a supplement, reported accurately were 57% (aOR = 0.57), while the odds that those with military service-connected insurance reported accurately were 19% (aOR = 0.19). With "no family history" of CRC in a first-degree relative as the reference group, the odds that those who said they had a first-degree relative with CRC reported accurately were 60% (aOR = 0.60). With those who had been patients between 1 and 24 years as the

Table 7. Final Multivariable Model of Main and Interaction Effects Associated with Accuracy, Defined as Concordance Between Self-report of Up-to-date Colonoscopy Screening Status and Documentation in Medical Record, Iowa Research Network (IRENE) Patients, 2010-2011.

Variable (n = 1,166)	aOR	Wald 95% CI	p-value
Education			
≥High School graduate	1.00	reference	
< High School graduate	0.40	(0.21-0.76)	0.0055
Insurance			0.0015
Group	1.00	reference	
Medicaid	0.93	(0.24-3.55)	0.9115
Medicare/Medicare + Supplement	0.57	(0.33-0.98)	0.0412
Military Service-connected	0.19	(0.08-0.45)	0.0002
Individual	0.79	(0.45-1.39)	0.4130
None	6.21	(0.82-47.04)	0.0774
Family History (first-degree relative) of Colorectal Cancer			
No	1.00	reference	
Yes	0.60	(0.39-0.91)	0.0154
Duration of Patient Status			0.0213
≥25 years	1.49	(0.97-2.29)	0.0677
1 - 24 years	1.00	reference	
< 1 year	0.22	(0.05-0.98)	0.0474
RUCA Category			0.0278
Recency of Last Visit			0.1038
Age			0.5048
RUCA Category x Age*			0.0345
Recency of Last Visit x Age*			0.0414

Abbreviations: CI = confidence interval; aOR = odds ratio, adjusted for all other variables in the model; RUCA = Rural-Urban Commuting Area.

* See Table 8 for an interpretation of the interaction effects.

reference group, the odds that those who had been patients for less than one year reported accurately were 22% (aOR = 0.22).

In the final model, Age significantly modified the association of both RUCA Category (p = 0.03) and Recency of Last Visit (p = 0.04) with accuracy when adjusted for all other variables in the model. Table 8 provides an interpretation of this effect modification for both variables at three selected ages. With those living in the RUCA category of "small town" as the reference group, at age 55, the odds that residents of the "metropolitan" and "large rural" categories reported accurately were 29% (aOR = 0.29), and 41% (aOR = 0.41), respectively, while the odds that those living in the "rural area" category reported accurately did not differ significantly from the reference group. At age 75, however, these differences disappeared, and the odds that those living in the "metropolitan", "large rural" and "rural area" categories reported accurately did not differ significantly from the reference group. For the Recency of Last Visit x Age interaction effect, with those who had visited their doctor within the past year as the reference group, the odds that those at age 65 who had seen their doctor over a year ago reported accurately were 44% (aOR = 0.44); at age 75, the odds that those whose last visit was greater than one year ago reported accurately were 25% (aOR = 0.25).

Table 8. Interpretation of Final Multivariable Model of Interaction Effects Associated with Accuracy, Defined as Concordance Between Self-report of Up-to-date Colonoscopy Screening Status and Documentation in Medical Record, at Three Ages, Iowa Research Network (IRENE) Patients, 2010-2011.

Variable (n = 1,166)	aOR	Wald 95% CI
RUCA Category x Age		
AT AGE 55		
Metropolitan	0.29	(0.12-0.69)
Large rural	0.41	(0.17-0.97)
Small town	1.00	reference
Rural area	0.49	(0.21-1.15)
AT AGE 65		
Metropolitan	0.90	(0.49-1.64)
Large rural	0.62	(0.37-1.05)
Small town	1.00	reference
Rural area	0.64	(0.38-1.08)
AT AGE 75		
Metropolitan	2.78	(0.86-8.99)
Large rural	0.94	(0.36-2.43)
Small town	1.00	reference
Rural area	0.82	(0.31-2.19)
Recency of Last Visit x Age		
AT AGE 55		
past year	1.00	reference
> past year	0.77	(0.44-1.34)
AT AGE 65		
past year	1.00	reference
> past year	0.44	(0.30-0.65)
AT AGE 75		
past year	1.00	reference
> past year	0.25	(0.12-0.54)

Abbreviations: CI = confidence interval; aOR = odds ratio, adjusted for all other variables in the model; RUCA = Rural-Urban Commuting Area.

DISCUSSION

Patient self-report of up-to-date CRC screening status was reasonably accurate in this sample based on estimates of sensitivity, specificity and concordance for "any" test and for CSPY and DCBE. Sensitivity was poor, however, for both FOBT/FIT and FSIG. Overreporting (based on R2R) was found for all tests, consistent with results from previous studies, and was substantially higher for those tests (FOBT, FSIG, DCBE) which had few documented results in the medical record. Almost all tests (95%) found in the MR were for screening by CSPY.

CSPY estimates were more accurate than other tests with regard to sensitivity, suggesting that, based on evidence in the medical record, patients knew that they were up-to-date with screening, and R2R, indicating less overreporting relative to other the tests. Consistent with the previous results of Shokar et al (2011), CSPY estimates were less accurate, however, when specificity and concordance were the estimates of interest. Results of this study suggest that when there was no evidence of screening by CSPY in the medical record, patients were less aware that they were *not* up-to-date. With MR as the reference standard and the awareness that such a standard is not always "gold" (Ferrante et al, 2008), it's also possible that documentation in the medical records was not complete or that oversights occurred during the abstraction process, and that, in fact, the patient *was* up-to-date. Though efforts were made secure evidence of CRC screening from outside offices if there was indication in the MR that CRC screening had been accomplished, it was not feasible to contact all subjects (about a quarter of all those without documentation) who suggested they were up-to-date with CRC screening but for whom no evidence was found in the MR.

Using a more liberal time interval to assess up-to-date status, e.g., within 11 rather than ten years for CSPY, made almost no difference when considering

up-to-date status for all tests, contrary to what was predicted. When indeterminate data were treated as missing and excluded from the analysis, accuracy decreased, most noticeably with regard to specificity estimates for "any" test. This is consistent with the findings of Vernon et al (2008), who suggested greater consequences for specificity estimates if there were a high proportion of uncertain and missing responses which were later excluded from analysis.

The design of the survey instrument used in this study, which assumed a physician recommendation prior to completion of a CRC screening test, contributed to the large number of indeterminate responses and the consequent decrease in specificity when such responses were excluded. If patients responded that they were unsure if their doctor had recommended a particular CRC test, for example, and then followed instructions to skip to the next set of questions related to a different CRC test, this did not allow them to directly answer the questions about whether they had actually had the test and, if so, if they had had it within the specified time frame. The impact of not being able to answer these two questions directly was magnified by the need to respond to such questions for each of the five CRC tests; thus, a participant's ability to respond definitively that they were *not* up-to-date with *any* of the five tests was greatly reduced, resulting in much lower estimates for specificity, e.g., specificity for "any" = 22/219 = 0.10 when indeterminate responses were not included, compared to 527/724 = 0.73 when such responses were considered nonadherent.

Efforts to minimize the effect of patients receiving treatment from several offices over time, potentially leading to less complete medical records, proved to be inconsequential. Restricting the analysis to only those patients who had been in IRENE offices for at least 10 years made very little difference in validity estimates.

Sociodemographic variables associated with accurate patient self-report in both univariate and multivariable models included education, insurance status and family history of CRC. Similar to the results of Partin et al (2008), the odds that those individuals who were at least high school graduates reported accurately that they were up-to-date with CRC screening by CSPY were higher than those with less education when adjusting for all other variables in the model. In accordance with previous studies, no association was found between gender or income and accuracy.

While family history of CRC in an immediate family member was associated with accuracy of self-report, the direction of the association was reversed from what has been found in previous studies (Partin et al, 2008). The results of this study found, in both univariate and multivariable analyses, that the odds for those with a family history of CRC to report accurately were 54% and 60%, respectively, when compared with the odds for those participants who did not have such a family history.

This result is somewhat counter-intuitive. One would expect that those who had a close family member with CRC would be more aware of and attentive to their own screening status. History of CRC was a self-reported variable; it's possible respondents weren't accurate in their reporting of such history because they were confusing the history of a first-degree relative who had CRC with one who actually had only a history of *polyps*. Similar to the misreporting resulting from cognitive issues related to testing (Vernon, Meissner et al, 2004), patients can have confusion over the difference between colorectal cancer and colorectal polyps.

Another significant association with self-report accuracy not previously found in the literature to date was insurance status. Adjusting for all other variables in the model, the odds for those individuals who were covered under

Medicare, with or without a supplemental policy, to report accurately were 57% when compared with the odds for those covered with group insurance through an employer; the odds for those with military service-connected insurance – albeit a small group in this study– to report accurately were only 19% when compared with the odds for those covered through their employment. These respondents were perhaps more likely to have CRC testing done at a VA facility, increasing the chances that corresponding results would not have been documented in the family medicine office's medical records.

Healthcare utilization factors were found to be associated with accurate self-report. Not surprisingly, the odds for those who had been patients in their family medicine offices for less than one year to report accurately were 22% when compared with those who had been patients from one to 24 years. One possible explanation for this is the time lag that can occur in obtaining medical records from previous offices when a patient changes from one primary care provider to another. In this scenario, the lower rate of concordance between patient self-report of up-to-date CSPY screening status and documentation in the medical record would be a consequence of this time lag and a reflection of incomplete data rather than misreporting on the part of the patient. The longevity of patients in their offices, with an average length of just over 17 years, *is* somewhat surprising, however, though perhaps not unique to those older individuals living in rural areas of the U.S.

Adjusting for other factors (Table 7) did not change the significant associations between sociodemographic and healthcare utilization variables and accuracy of patient self-report found in the univariate analysis (Table 6); age, however, did modify this relationship, at least with regard to recency of last patient visit to the IRENE office and RUCA category. At younger ages, there was no significant difference in the odds for those patients who saw their doctor more

than one year ago reporting accurately about CSPY when compared with those who did see their doctor within the past year. However, at older ages, this relationship changed: the odds for those who hadn't been in for an exam within a year to report accurately were 44% at age 65 and 25% at age 75 when compared with those patients who *had* had an annual visit.

This result underscores the importance of regular visits to the doctor for older individuals. Less contact with one's physician may be associated with less accurate reporting which may act as a proxy for an individual's personal health awareness or level of health literacy. The question of the relationship of patient-physician communication and accuracy of patient self-report of CRC testing is beyond the scope of the present study, but evidence does suggest that more comprehensive discussion of CRC screening by one's primary care physician is associated with increased CRC screening by patients, independent of healthcare utilization factors (Levy et al, 2006; Mosen et al, 2013).

When adjusted for all other variables, age in this study also modified the rural-urban residence, i.e., RUCA, effect. At younger ages, e.g., age 55, the odds for those living in areas classified as "large rural" and "metropolitan" to report accurately were 59% and 71% lower, respectively, compared with those living in a "small town". At older ages, however, there were no significant differences in the odds that one would report accurately based on residence according to RUCA category classification.

This result is interesting and not wholly in line with what might be expected. In this study, the rural-urban continuum with regard to accuracy of self-report of CSPY does not follow a simple ordinal pattern. The odds for patients who reside in more rural settings to report accurately are not increasingly lower when compared with those who live in more urban settings, as was initially hypothesized, and age is also a contributing factor. At younger

ages, a patient's place of residence plays a significant role with regard to the odds of accurate self-report of up-to-date status of colorectal cancer screening, a significance that disappears at older ages.

Reasons for this difference are unclear. One might speculate that younger small town and rural residents may have a different and closer relationship with their doctor, as reflected in higher odds of being more accurate reporters of up-to-date CRC screening status, which the younger urban and suburban dweller doesn't have or need to have. At older ages, however, the benefits of a closer relationship with one's health care provider extend to all, regardless of where one lives. It's also possible that younger residents of urban and suburban areas, with their greater potential access to specialists, at least in terms of geographic proximity, might be more likely to obtain CRC screening without involving or sharing the results with their primary care provider. Incomplete medical records in the primary care provider's office might explain the lower odds that patients living in these geographic locales reported accurately, a scenario that perhaps is less likely to occur with older patients who have more health problems and consequent need for more diligent and vigilant following by their physician.

Previous studies suggest that rural residents, compared to their more urban counterparts, may face numerous barriers to obtaining preventive health services and screening for early detection of cancers (Levy et al, 2006, Coughlin et al, 2004, McLafferty and Wang, 2009). These barriers include poor geographic access to primary care (Jones et al, 2009) and cancer screening services, lack of insurance (apparently not an issue in this study's sample) and lack of knowledge regarding screening guidelines. Though the odds for those with higher education to report accurately that they were up-to-date with CRC screening by CSPY were higher than those with less education, specific knowledge with regard to screening guidelines was not directly assessed in this project.

McLafferty and Wang (2009) elaborate that the concentration of vulnerable population groups, including the elderly, in rural areas increases the significance of the barriers described above. Rural residents are more likely than their urban counterparts to delay seeking medical care or may not seek it altogether (Jones et al, 2009). Peek-Asa et al (2011) describe the physical and social isolation of rural residents and, along with Coughlin et al (2004), call for increased focus on preventive services for rural populations.

The results of this study suggest that more resources may also be needed for those younger residents living in metropolitan areas and the "large rural" towns surrounding them. In Iowa, with its changing demographics, the way rurality is measured has increased significance. Over the period from 2000 to 2010, the state's population grew increasingly urban, with seven counties – all near urban centers – growing by more than ten percent, while five counties in rural western Iowa lost at least ten percent of their residents (U.S. Bureau of the Census, retrieved April 27, 2014).

Results from the present study confirm that the designation of rurality in Iowa by county may not be the best unit of measurement when investigating spatial factors. Earlier support for this can be found in a report on colorectal cancer screening in Iowa by Thompson et al (2006); in it, these authors conclude that it would be misleading for rural communities located within more densely populated counties to receive a designation of the entire county as urban. Indeed, the particular rurality categorization scheme used may result in very different conclusions. Previous studies (Coughlin et al, 2004; Anderson et al, 2013) exploring the influence of geographic factors on utilization of CRC screening have made use of a scheme involving only two or three rural-urban categories, either grouping the more isolated rural areas with small towns or collapsing them into a potentially overly-simplistic rural-urban dichotomy.

Research exploring health-related issues, including the accuracy of patient self-report of CRC screening, needs to take into account the increasingly-relevant geographic variable as a correlate, tailoring its classification accordingly. The RUCA classification scheme, a collaborative project between the federal Health Resources and Service Administration's (HRSA) Office of Rural Health Policy, the U.S. Department of Agriculture's Economic Research Service and the Washington, Wyoming, Alaska, Montana and Idaho (WWAMI) Rural Health Research Center, is most versatile in its applicability, but thoughtful consideration must be given to the particular way codes are aggregated which is dependent on one's objectives. The more discriminating four-group Categorization A (WWAMI Rural Health Research Center, retrieved April 27, 2014) used in the present study should be considered for regions with substantial rural populations to avoid the masking of particular effects in prevalence estimates.

Strengths

A unique aspect of this validity study is its "real world" applicability: its participants were drawn from a practice-based research network where they reside in geographically-diverse areas of varying population densities within a primarily rural state. The use of the RUCA classification scheme allowed for a discriminating classification of rurality to account for the range of geographic variability in a participant's place of residence. Additional strengths of the current study include use of a survey instrument with detailed CRC test descriptions and defined time intervals, a high percentage (88%) of medical records reviewed in a largely stable population and a comprehensive assessment of potential measurement error in validity estimates, including attempts to account for lack of completeness of available medical records by tracking down and requesting evidence of reported CRC screenings at facilities outside of the

IRENE office and conducting a separate analysis for those patients who had been in their particular IRENE office for at least ten years.

Limitations

Potential limitations of this study include selection bias, in that participants had to give written consent to participate, resulting in a more willing, interested group; instrument bias, affecting participants' reports of CRC screening, as questions were linked to physician recommendation, resulting in a greater number of indeterminate responses; lack of patient cognitive assessment, which could have had an effect on a patient's ability to report accurately; and potential lack of generalizability to other areas in the U.S. since participants were from a primarily rural, white, non-Hispanic Midwestern state.

The large number of indeterminate responses did not have as great an effect in this project, with its focus on individuals who had been screened for CRC cancer, as it might have. As Vernon et al (2008) identified, indeterminate responses would have the greatest effect on estimates of specificity and would have less importance for obtaining valid estimates of sensitivity. In addition, the choice of concordance as the measure of validity in the secondary analyses proved to be a good one, as results from the primary analyses showed this particular estimate to be essentially unchanged depending on whether indeterminate responses were excluded or considered nonadherent.

Future Directions

Recommendations for future research center around methodological standardization: use of the validated survey proposed by Vernon et al (2008) to enhance comparability in survey instrument operational definitions and measurement construction as well as outcome measures when assessing correlates of CRC self-report accuracy; choice of appropriate geographic classification, i.e., consideration of "place" as an independent variable and choice

of the RUCA classification scheme which appropriately reflects the geographic area to be studied; investigation of the association of physician recommendation, patient health status and other patient-provider relationship variables with self-report accuracy; exploration of validity estimates for developing technologies, e.g., computed tomographic colonography and stool DNA tests; and comparison of validity estimates using paper vs. electronic medical records. Recommendations for providers include the need to go after medical records from other offices where patients have been seen previously as well as records from specialists', i.e., gastroenterologists', offices and the hospitals where they may practice.

Patient self-report of up-to-date CRC screening status will continue as a singular and important source of data for population-based estimates of the prevalence of CRC screening in the U.S. Results from this study suggest that patient self-report, while generally accurate, is influenced by a number of sociodemographic characteristics as well as healthcare utilization factors related to the patient. It therefore remains incumbent upon practitioners and researchers alike to be aware of and make every attempt to control for as many sources of variability in validity estimates as possible in order to better serve the health of the nation.

APPENDIX A

LITERATURE REVIEWS

Table A1. Literature Review of Primary Validity Analysis Studies of Colorectal Cancer Screening Self-report with Medical Record as Reference Standard

Author (Year)	Subjects	Recency	Sens	Spec	Concord	R2R
Baier et al (2000)	40-75 years Managed Care CO (urb) White	Ever	FOBT = 0.96 (n = 186) CSPY = 0.89 (n = 97) FSIG = 0.95 (n = 79)	FOBT = 0.86 (n = 64) CSPY = 0.97 (n = 95) FSIG = 0.92 (n = 90)		
Schoen et al (2002)	50-79 years Medic recip/lic driv PA (suburban) White	Past 5 years	Any = n/a CSPY = n/a FSIG = n/a DCBE = n/a			

Table A1. Continued

Lipkus et al (2003)	50 - 75 years Union fund NJ White/male	Past year	FOBT = 0.69 (n = 445)	FOBT = 0.85 (n = 213)	FOBT= 0.74 (n = 658)	FOBT = 0.76 (n = 445)
Madlensky et al (2003)	35+ years 1st deg rel of CRC pts Ontario (Canada)	Ever	FOBT = 0.92 (n = 37) CSPY = 0.95 (n = 164) FSIG = 0.79 (n = 24)	FOBT= 0.70 (n= 269) CSPY = 0.92 (n = 168) FSIG = 0.80 (n = 299)	FOBT = 0.72 (n = 306) CSPY = 0.94 (n = 332) FSIG = 0.80 (n = 323)	FOBT = 3.14 (n = 37) CSPY = 1.03 (n = 164) FSIG = 3.29 (n = 24)
Hall et al (2004)	Men 45+; Women 55+ Managed care CA/GA/MN	Past two years (FOBT) Past five years (CSPY/ FSIG)				FOBT = 1.92 (n = 251) CSPY = 1.47 (n = 211) FSIG = 1.38 (n=678)

Table A1. Continued

Fiscella et al (2006)	65+ years Medic recip U.S. White/urb/supp insur	Past year				Any = 1.29 (n = 432)
Bastani et al (2008)	40 - 80 years 1st deg rel of CRC pts CA/U.S. AA/Hispanic/Asian	Past year	Any = 0.93 (n = 42) FOBT = 0.73 (n = 15) CSPY = 0.84 (n = 25) FSIG = 0.83 (n = 6)	Any = 0.64 (n = 81) FOBT = 0.75 (n = 108) CSPY = 0.83 (n = 98) FSIG = 0.97 (n = 117)	Any = 0.74 (n = 123) FOBT = 0.75 (n = 123) CSPY = 0.83 (n = 123) FSIG = 0.96 (n = 123)	Any = 1.62 (n = 42) FOBT = 2.53 (n = 15) CSPY = 1.52 (n = 25) FSIG = 1.50 (n = 6)
Ferrante et al (2008)	50+ years Comm primary care NJ White/female/ins	Past year	FOBT = 0.51 (n = 90)	FOBT = 0.79 (n = 633)	FOBT = 0.76 (n = 723)	FOBT = 1.96 (n = 90)

Table A1. Continued

Jones et al (2008)	50+ years Comm/lic driv MN (5 counties) White/rural	Past year (FOBT) Past 5 years (FSIG/ DCBE) Past 10 years (CSPY)	Any = 0.96 (n = ?) FOBT = 0.93 (n = ?) CSPY = 1.00 (n = ?) FSIG = 0.87 (n = ?) DCBE = 0.74 (n = ?)	Any = 0.81 (n = ?) FOBT = 0.90 (n = ?) CSPY = 0.87 (n = ?) FSIG = 0.87 (n = ?) DCBE = 0.93 (n = ?)		
Partin et al (2008)	50-75 years VAMC patients MN White/male/low inc	Past year (FOBT) Past 5 years (FSIG/ DCBE) Past 10 years (CSPY)	Any = 0.98 (n = 233) FOBT = 0.82 (n = 62) CSPY = 0.97 (n = 124) FSIG = 0.75 (n = 92) DCBE = 0.63 (n = 8)	Any = 0.59 (n = 87) FOBT = 0.89 (n = 253) CSPY = 0.72 (n = 199) FSIG = 0.76 (n = 222) DCBE = 0.85 (n= 296)	Any = 0.88 (n = 320) FOBT = 0.87 (n = 320) CSPY = 0.81 (n = 320) FSIG = 0.76 (n = 320) DCBE = 0.85 (n = 320)	Any = 1.14 (n = 233) FOBT = 1.31 (n = 62) CSPY = 1.42 (n = 124) FSIG = 1.33 (n = 92) DCBE = 6.13 (n = 8)
Powe et al (2008)	50+ years 4 FQHCs GA AA/female/low inc	Past year	FOBT = 0.50 (n = 4)	FOBT = 0.50 (n = 12)	FOBT = 0.50 (n = 16)	FOBT = 2.00 (n = 4)

Table A1. Continued

Schenck et al (2008)	50 - 80 years Medic recip NC (10 counties) White/AA/urb	Past year	FOBT = 0.47 (n = 109)	FOBT = 0.76 (n = 452)	FOBT = 0.70 (n = 561)	FOBT = 1.48 (n = 109)
Vernon et al (2008)	51 - 74 years Multispecialty clinic TX urb/semi-rural White/AA/Hispanic	Past year (FOBT) Past 5 years (CSPY/ FSIG/DCBE)	FOBT = 0.82 (n = 138) CSPY = 0.91 (n = 232) FSIG = 0.76 (n = 219) DCBE = 0.56 (n = 103)	FOBT = 0.86 (n = 719) CSPY = 0.91 (n = 625) FSIG = 0.89 (n = 638) DCBE = 0.97 (n = 754)	FOBT = 0.85 (n = 857) CSPY = 0.91 (n = 857) FSIG = 0.85 (n = 857) DCBE = 0.92 (n = 857)	FOBT = 1.57 (n = 138) CSPY = 1.15 (n = 232) FISG = 1.10 (n = 219) DCBE = 0.82 (n = 103)
Shokar et al (2011)	50 - 80 years Univ primary care TX urb/semi-rural White/AA/Hispanic	Past year (FOBT) Past 5 years (FSIG/ DCBE) Past 10 years (CSPY)	Any = 0.85 (n = ?) FOBT = 0.70 (n = 25) CSPY = 0.89 (n = 93) FSIG = 0.50 (n = 22) DCBE = 0.45 (n = ?)	Any = 0.69 (n = ?) FOBT = 0.89 (n = 246) CSPY = 0.86 (n = 178) FSIG = 0.91 (n = 249) DCBE = 0.83 (n = ?)	Any = 0.76 (n = ?) FOBT = 0.87 (n = 271) CSPY = 0.87 (n = 271) FSIG = 0.87 (n = 271) DCBE = 0.81 (n = ?)	Any = 1.25 (n = ?) FOBT = 1.80 (n = 25) CSPY = 1.16 (n = 93) FSIG = 1.55 (n = 22) DCBE = 4.55 (n = ?)

Table A1. Continued

Reiter et al (2013)	51 - 75 years Random comm sample OH (12 counties) White/rural	Past year (FOBT) Past 5 years (FSIG/ DCBE) Past 10 years (CSPY)	Any = 0.95 (n = 356) FOBT = 0.32 (n = 31) CSPY = 0.96 (n = 340) FSIG = 0.17 (n = 6)	Any = 0.58 (n = 365) FOBT = 0.93 (n = 690) CSPY = 0.65 (n = 381) FSIG = 0.96 (n = 715)	Any = 0.76 (n = 721) FOBT = 0.90 (n = 721) CSPY = 0.80 (n = 721) FSIG = 0.96 (n = 721)	Any = 1.38 (n = 356) FOBT = 1.87 (n = 31) CSPY = 1.35 (n = 340) FSIG = 4.50 (n = 6)
Beebe et al (2014)	49 - 85 years Random comm sample MN (1 county) White/urb/rural	Ever	Any = 0.95 (n = 2,929) FOBT = 0.38 (n = 295) CSPY = 0.86 (n = 2,754) FSIG = 0.63 (n = 1,604) DCBE = 0.92 (n = 76)	Any = 0.65 (n =709) FOBT = 0.80 (n = 3,343) CSPY = 0.81 (n = 884) FSIG = 0.86 (n = 2,034) DCBE = 0.72 (n = 3,562)	Any = 0.89 (n = 3,638) FOBT = 0.76 (n = 3,638) CSPY = 0.85 (n = 3,638) FSIG = 0.76 (n = 3,638) DCBE = 0.72 (n = 3,638)	Any = 1.03 (n = 2,929) FOBT = 2.66 (n = 295) CSPY = 0.92 (n = 2,754) FSIG = 0.80 (n = 1,604) DCBE = 22.30 (n = 76)

Abbreviations:
Validity Estimates: Sens = sensitivity (True Positive / True Positive + False Negative); Spec = specificity (True Negative / True Negative + False Positive); Concord = concordance (True Positive + True Negative / total sample); R2R =r eport-to-records ratio (True Positive + False Positive / True Positive + False Negative).
Screening Tests: FOBT = fecal occult blood test; CSPY = colonoscopy; FSIG = flexible sigmoidoscopy; DCBE = double-contrast barium enema.
States: CA = California; CO = Colorado; GA = Georgia; MN = Minnesota; NC = North Carolina; NJ = New Jersey; OH = Ohio; PA = Pennsylvania; TX = Texas.
Miscellaneous: AA = African-American; comm = community-based; CRC = colorectal cancer; FQHCs = federally-qualified health centers; 1st deg rel = first-degree relatives; inc = income; ins = insured; lic driv = licensed drivers; Medic recip = Medicare recipients; n/a=validity estimate other than Sens/Spec/Conc/R2R provided ; ? = not provided/unable to be determined; pts= patients; supp insur = supplemental insurance; univ = university-based; urb = urban; VAMC = veterans affairs medical center.

Table A2. Literature Review of Secondary Validity Analysis Studies of Colorectal Cancer Screening Self-report with Medical Record as Reference Standard

Author (Year)	Subjects	Dep Var	Ind Var (Ref Group)	Test	Estimate	Signif P-value
Baier et al (2000)	40 – 75 yrs Managed Care CO (urb) White	Sens/Spec	**Gender**[1] **Race/Ethnicity** Age Fam Hx CRC	FSIG FOBT	Spec = 0.98 (female) > 0.83 (male) Sen s = 0.98 (White) > 0.89 (nonwhite)	0.004 0.03
Lipkus et al (2003)	50 – 75 yrs Union fund NJ White/male	Conc	**Age** (ref = 50 years) Fam Hx CRC "other demo" variable	FOBT	OR = 1.03 (1 year increments)	0.03
Hall et al (2004)	Men 45+/ Women 55+ Managed Care CA/GA/MN	R2R	**Gender/Race** (ref = AA males) **Age** (ref = 40-49 yrs) Edu	CSPY FOBT FSIG CSPY	OR = 0.51 (White/other males) OR = 0.44 (Females) OR = 2.78 (70+ years) OR = 2.85 (70+ years) OR = 14.36 (70+ years)	0.05 0.05 0.05 0.05 0.05
Fiscella et al (2006)	65+ yrs Medic recip U.S. White/urb/ Supp insur	FP	Gender **Race** (ref = White) Age Edu Inc Supp Insur Health status Proxy response	Any	OR = 1.92 (nonwhite) aOR = 1.68 (nonwhite)	0.05 0.05

Table A2. Continued

Partin et al (2008)	50 – 75 yrs VAMC pts MN White/Male/ Low inc	Sens/Spec/ Conc/R2R	Gender/Race **Age**	Any	Sens = 1.00 (65 -75 yrs) > 0.97 (50 - 64 yrs)	0.05
					Spec = 0.45 (65 - 75 yrs) < 0.67 (50 - 64 yrs)	0.05
					Conc = 0.84 (65 - 75 yrs) < 0.89 (50 - 64 yrs)	
					R2R = 1.22 (65 - 75 yrs) > 1.09 (50 - 64 yrs)	0.05
			Edu	Any	Sens = 1.00 (Coll grad) = 1.00 (HS diploma)	
					Spec = 0.79 (Coll grad) = 0.46 (HS diploma)	0.05
					Conc = 0.95 (Coll grad) > 0.85 (HS diploma)	0.05
					R2R = 1.06 (Coll grad) < 1.21 (HS diploma)	0.05
			Inc/Mar Stat			
			Fam Hx CRC	Any	Sens = 0.98 (Fam hx) < 0.99 (No fam hx)	
					Spec = 0.43 (Fam hx) < 0.58 (No fam hx)	
					Conc = 0.93 (Fam hx) > 0.86 (No fam hx)	
					R2R = 1.05 (Fam hx) < 1.19 (No Fam hx)	0.05
			Comorb/Psych dx			
Griffin et al (2009)	50 – 75 yrs VAMC MN White/Male/ Low inc	Conc/R2R	**Gender**	Any	Conc = 0.89 (Females) > 0.85 (Males)	
				FOBT	Conc = 0.88 (Females) < 0.89 (Males)	
				CSPY	Conc = 0.89 (Females) > 0.76 (Males)	0.01
				FSIG	Conc = 0.83 (Females) > 0.69 (Males)	0.01
				DCBE	Conc = 0.91 (Females) > 0.84 (Males)	
					Any R2R = 1.03 (Females) < 1.09 (Males)	
					FOBT R2R = 1.21 (Females) > 1.11 (Males)	
					CSPY R2R = 1.21 (Females) < 1.45 (Males)	
					FSIG` R2R = 1.17 (Females) > 1.10 (Males)	
					DCBE R2R = 2.00 (Females) < 6.17 (Males)	
			Race/Age/Edu/Mar Stat/ Comorb/Psych dx/Recom			

Table A2. Continued

Shokar et al (2011)	50 – 80 yrs Univ prim care TX urb/ semi-rural	Sens/Spec/ Conc/R2R	Race/Ethnicity	Any	Sen s= 0.77 (Hisp) < 0.83 (AA) < 0.93 (White)	0.05
				FOBT	Sens = 0.44 (Hisp) < 0.78 (AA) < 0.89 (White)	0.05
				CSPY	Sens = 0.81 (AA) < 0.86 (Hisp) < 1.00 (White)	
				FSIG	Sens = 0.25 (Hisp) < 0.50 (White) < 0.60 (AA)	0.05
				DCBE	Sens = 0.25 (Hisp) < 0.50 (AA) < 0.67 (White)	0.05
White et al (2013)	51 – 74 yrs Multi-specialty clinic TX urb/ semi-rural White/AA/Hisp	Sens/Spec/ Conc/R2R	Gender /Race/Ethnicity/ Age/Edu/Mar Stat/Fam Hx CRC/# visits past 5 yrs			
			# visits out office	Any	Sens = 0.88 (6+ visits) < 0.91 (0-5 visits)	
					Spec = 0.56 (6+ visits) < 0.77 (0-5 visits)	0.05
					Conc = 0.69 (6+ visits) < 0.85 (0-5 visits)	0.05
					R2R = 1.51 (6+ visits) > 1.10 (0-5 visits)	
			Prov recom test	Any	Sens = 0.93 (prov rec) > 0.56 (no prov rec)	0.05
					Spec = 0.66 (prov rec) < 0.95 (no prov rec)	0.05
					Conc = 0.82 (prov rec) < 0.88 (no prov rec)	
					R2R = 1.16 (prov rec) > 0.83 (no prov rec)	
Reiter et al (2013)	51 – 75 yrs Random comm sample MN (1 county) White/urb/rural	PPV	Gender/Race/Age Edu/Inc/Mar Stat/Ins/			
			Ann Exam in 2 yrs	Any	OR = 2.58 (“yes”)	0.05
			(re f = ”no”)		aOR = 2.78 (“yes”)	0.05
			Self-rated health	Any	OR = 1.80 (“good/very good/excellent”)	0.05
			(ref = ”poor/fair”)		aOR = 1.88 (“good/very good/excellent”)	0.05
			Employ/Smoking/ Med cond/Risk/ Worry/Aware test			

Table A2. Continued

---

[1] Significant independent variables are **bolded**.

Abbreviations:
Validity Estimates: Concord = concordance (True Positive + True Negative / total sample); R2R = report-to-records (True Positive + False Positive / True Positive + False Negative); Sens = sensitivity (True Positive / True Positive + False Negative); Spec = specificity (True Negative / True Negative + False Positive).
Screening Tests: FOBT = fecal occult blood test; CSPY = colonoscopy; FSIG = flexible sigmoidoscopy; DCBE = double-contrast barium enema.
States: CA = California; CO = Colorado; GA = Georgia; MN = Minnesota; NJ = New Jersey; TX = Texas.
Statistical: aOR = adjusted odds ratio; FP = false positives; OR = odds ratio; ref = reference; PPV = positive predictive value; signif = significant.
Miscellaneous: AA = African-American; ann = annual; coll = college; comm = community-based; comorb = comorbidity; CRC = colorectal cancer; demo = demographic; edu = education; employ = employed; fam hx = family history; grad = graduate; HS = high school; Hisp = Hispanic; inc = income; ins = insured; mar stat = marital status; med cond = medical condition; Medic recip = Medicare recipients; prim = primary; prov = provider; psych dx = psychiatric diagnosis; recom = recommendation; supp insur = supplemental insurance; univ = university-based; urb = urban; VAMC = Veterans Affairs Medical Center; yrs = years; # = number.

APPENDIX B
INFORMED CONSENT DOCUMENT

INFORMED CONSENT DOCUMENT

Project Title: **Interventions to Improve Colon Cancer Screening in Rural Iowa Counties**

Research Team: **Barcey Levy, MD, PHD, John Ely, MD, MSPH Jeanette Daly, PHD, RN, Mary Merchant, PHD, RN, Joyce Baker, MSW**

This consent form describes the research study to help you decide if you want to participate. This form provides important information about what you will be asked to do during the study, about the risks and benefits of the study, and about your rights as a research subject.

- If you have any questions or do not understand something in this form, you should ask the research team for more information.
- You should discuss your participation with anyone you choose such as family or friends.
- Do not agree to participate in this study unless the research team has answered your questions and you decide that you want to be part of this study.

WHAT IS THE PURPOSE OF THIS STUDY?

This is a research study. We are inviting you to participate in this research study because you are a healthy individual between the ages of 52 and 79 years who attends a family practice office in the state of Iowa. Your family physician's office is a member of the Iowa Research Network (IRENE). IRENE is a practice-based research network. IRENE's key mission is to create new knowledge with relevance to rural primary care clinicians and their patients, with the outcome of improving the care of patients.

The purpose of this research study is to test office reminder systems to improve colon cancer screening. Colon cancer is a disease that is largely preventable or very treatable if detected early. However, only about half of the eligible U.S. population has been screened for this disease.

Early screening is important because:

- Colon cancer starts with NO SYMPTOMS in people who feel perfectly fine.
- It is more common in people after age 50.
- It begins with a specific type of polyp, which can be detected and removed during a colonoscopy. Once removed, these polyps pose no threat to your health. (There are other types of colon polyps that cause no health problems.)

- Sometimes these polyps bleed, but very sensitive tests are needed to detect the bleeding.
- The people of Iowa have one of the highest rates of colon cancer in the nation.
- Colon cancer is a HIGHLY TREATABLE AND CURABLE disease if caught early.
- It is estimated that close to 90% of colon cancers could be prevented with appropriate screening.

Colon cancer screening is part of high-quality medical care and is recommended regularly after age 50 by all major medical organizations that develop preventive guidelines for adults.

HOW MANY PEOPLE WILL PARTICIPATE?

Approximately 2,120 people will take part in this study conducted by investigators at the University of Iowa.

HOW LONG WILL I BE IN THIS STUDY?

If you agree to take part in this study, your involvement will last for about 1 year as described below. The total time you will spend participating will be no more than 15 minutes per survey for a maximum of three surveys. Some subjects will be asked to watch a DVD that has a program that is 8 minutes in length.

WHAT WILL HAPPEN DURING THIS STUDY?

If you agree to participate, you will be randomized to one of 4 groups. This means that you will be assigned to one of 4 groups based on a procedure similar to the flip of a coin. You will have an equal chance of being in one of the four groups. We will test strategies of increasing intensity to see whether individuals will be tested for colon cancer. These strategies will include things like reminding your physician to order this test and mailing educational materials to you (both written and in DVD format) to educate you about CRC screening. Some individuals will receive a telephone reminder.

At the conclusion of the study, individuals who were not randomly assigned to the patient education groups will receive a mailing including this education.

We will ask you to complete the following surveys which will each take no more than 10 minutes of your time:

- A baseline survey (enclosed)
- A survey after mailed educational materials (if you are in the educational materials group)
- A 1-year follow-up survey.

You are free to skip any questions you do not wish to answer. If some of your answers are not clear, a member of the research team may call you for

clarification.

HIPAA INFORMATION

In about 1 year, we will review your medical record for dates and types of colon cancer screening tests, medical diagnoses and medications used. We will link this information with your survey responses. Your information will be identified by a unique number and not with your name or other identifying information.

WHAT ARE THE RISKS OF THIS STUDY?

You may experience one or more of the risks indicated below from being in this study. In addition to these, there may be other unknown risks, or risks that we did not anticipate, associated with being in this study.

During this study, you may receive information about screening tests for colon cancer and be encouraged to undergo screening. Screening is recommended for all average-risk individuals over the age of 50 years to help prevent colon cancer by all major organizations who write preventive guidelines for physicians. The risks of screening, if you undergo it through this study, are no different than if you were to undergo these tests outside of this study. Some of the questions may make you feel uncomfortable.

Some individuals will be randomized to receive a test for fecal occult blood. Results of this test will be provided to you and your physician's office. If you choose to return this test, it is possible the test will indicate that you have blood in your stool, which may cause you to worry.

WHAT ARE THE BENEFITS OF THIS STUDY?

We don't know if you will benefit from being in this study. However, we hope that, in the future, other people might benefit from this study because we may have a better understanding of what strategies work to increase screening in primary care.

WILL IT COST ME ANYTHING TO BE IN THIS STUDY?

You will not have any costs for being in this research study. The fecal occult blood test received through this study will be developed at no cost to you. You and/or your medical/hospital insurance carrier will remain responsible for your regular medical care expenses, including any follow-up tests or appointments that might be recommended based on the results of the fecal occult blood test.

WILL I BE PAID FOR PARTICIPATING?

You will be paid for being in this research study. You will need to provide your address if a check will be mailed to you.

The $2 enclosed with this research mailing is yours to keep whether or not you decide to participate.

You will receive :

- $20 for completion of the Baseline Patient Survey and a signed Informed Consent.
- $10 for completion of the 1-year Follow-up Survey.
- Subjects in groups that receive educational materials will receive $10 upon completion of a follow-up survey.

WHO IS FUNDING THIS STUDY?

The American Cancer Society is funding this research study. This means that the University of Iowa is receiving payments from the American Cancer Society to support the activities that are required to conduct the study. No one on the research team will receive a direct payment or increase in salary from the American Cancer Society for conducting this study.

WHAT ABOUT CONFIDENTIALITY?

We will keep your participation in this research study confidential to the extent permitted by law. However, it is possible that other people such as those indicated below may become aware of your participation in this study and may inspect and copy records pertaining to this research. Some of these records could contain information that personally identifies you.

- federal government regulatory agencies,
- American Cancer Society
- auditing departments of the University of Iowa, and
- the University of Iowa Institutional Review Board (a committee that reviews and approves research studies)

To help protect your confidentiality, we will store all information related to the study in locked filing cabinets and/or password protected computers files. Your information will be linked with an identification number and not with your name. If we write a report or article about this study, the information will be summarized in such a way that no individuals can be identified.

WILL MY HEALTH INFORMATION BE USED DURING THIS STUDY?

The Federal Health Insurance Portability and Accountability Act (HIPAA) requires your health care provider to obtain your permission for the research team to access or create “protected health information” about you for purposes of this research study. Protected health information is information that personally identifies you and relates to your past, present, or future physical or mental health condition or care. We will access or create health information about you, as described in this document, for purposes of this research study. Once your health care provider has disclosed your protected health information to us, it may no longer be protected by the Federal HIPAA privacy regulations, but we will continue to protect your confidentiality as described under “Confidentiality.”

We may share your health information related to this study with other parties including federal government regulatory agencies, the American Cancer Society, and the University of Iowa Institutional Review Boards and support staff.

You cannot participate in this study unless you permit us to use your protected health information. If you choose *not* to allow us to use your protected health information, we will discuss any non-research alternatives available to you. Your decision will not affect your right to medical care that is not research-related. Your signature on this Consent Document authorizes your health care provider to give us permission to use or create health information about you.

Although you may not be allowed to see study information until after this study is over, you may be given access to your health care records by contacting your health care provider. Your permission for us to access or create protected health information about you for purposes of this study has no expiration date. You may withdraw your permission for us to use your health information for this research study by sending a written notice to Barcey T. Levy, 01292 E Pomerantz Family Pavilion, Department of Family Medicine, University of Iowa, 200 Hawkins Drive, Iowa City, Iowa 52242. However, we may still use your health information that was collected before withdrawing your permission. Also, if we have sent your health information to a third party, such as the study sponsor, or we have removed your identifying information, it may not be possible to prevent its future use. You will receive a copy of this signed document.

IS BEING IN THIS STUDY VOLUNTARY?

Taking part in this research study is completely voluntary. You may choose not to take part at all. If you decide to be in this study, you may stop participating at any time. If you decide not to be in this study, or if you stop participating at any time, you won't be penalized or lose any benefits for which you otherwise qualify.

Will I Receive New Information About the Study while Participating?

If we obtain any new information during this study that might affect your willingness to continue participating in the study, we'll promptly provide you with that information.

WHAT IF I HAVE QUESTIONS?

We encourage you to ask questions. If you have any questions about the research study itself or if you experience a research related injury, please contact Joyce Baker toll-free at 1-866-890-5963.

If you have questions, concerns, or complaints about your rights as a research subject or about research related injury, please contact the Human Subjects Office, 340 College of Medicine Administration Building, The University of Iowa, Iowa City, Iowa, 52242, (319) 335-6564, or e-mail irb@uiowa.edu. General information about being a research subject can be found by clicking "Info for Public" on the Human Subjects Office web site, http://research.uiowa.edu/hso. To offer input about your experiences as a research subject or to speak to someone other than the research staff, call the Human Subjects Office at the number above.

This Informed Consent Document is not a contract. It is a written explanation of what will happen during the study if you decide to participate. You are not waiving any legal rights by signing this Informed Consent Document. Your signature indicates that this research study has been explained to you, that your questions have been answered, and that you agree to take part in this study. You will receive a copy of this form.

We need your name both <u>printed below and then signed in the box</u>.

Subject's Name (print your name below):

_____

Do not sign this form if today's date is on or after $STAMP_EXP_DT.

_____

(Signature of Subject) (Date)

APPENDIX C
BASELINE SURVEY (EXCERPTED)

COLORECTAL CANCER
IOWA RESEARCH NETWORK STUDY
AMERICAN CANCER SOCIETY

BASELINE PATIENT SURVEY

Colon cancer is cancer of the large intestine or rectum.

1. What is your gender? ☐$_1$ Male ☐$_0$ Female

2. My primary health care provider is: ☐$_1$ Male ☐$_0$ Female

3. Do you have a **personal history** of colon cancer?
   ☐$_1$ Yes ☐$_0$ No
   ☐$_9$ Unsure

4. Do you have an **immediate** family member who has or had colon cancer? (mother, father, sister, brother)
   ☐$_1$ Yes ☐$_0$ No
   ☐$_9$ Unsure

5. Do you have a **more distant relative** who has or had colon cancer? (grandparent, cousin, aunt or uncle)
   ☐$_1$ Yes ☐$_0$ No
   ☐$_9$ Unsure

6. Do you have a personal history of **ulcerative colitis** or **Crohn's disease**?
   ☐$_1$ Yes ☐$_0$ No
   ☐$_9$ Unsure

Colon cancer screening is done to detect possible problems **before** you have symptoms. There are several methods available for colon cancer screening, including:

- The take home fecal occult blood test (FOBT) where samples of 3 consecutive stools are placed on cards and the cards are brought back to your doctor to be checked for blood

- The take home fecal immunochemical test (FIT) another type of FOBT that also uses cards to collect stool samples and is considered more accurate than other types, with fewer restrictions for the patient.
- The flexible sigmoidoscopy where the doctor uses a lighted instrument to examine about one-third of your colon (usually done in the doctor's office without sedation)
- The colonoscopy where your entire colon is examined using a flexible instrument with a light on it (you are usually given medication to help with pain and make it more comfortable; someone needs to drive you home afterwards)
- Barium enema where the colon is filled with liquid barium and an x-ray is taken

7. Has a doctor or nurse **ever** talked with you about having a test for colon cancer? ☐$_1$ Yes ☐$_0$ No ☐$_9$ Unsure

8. Has your physician ever recommended you undergo colon cancer screening? (Screening is to detect possible problems before you have symptoms.) ☐$_1$ Yes ☐$_0$ No(go to question 9) ☐$_9$ Unsure

   8a. Approximately how old were you when this recommendation was **first** made?
   _____ years old

   8b. Approximately how old were you when this recommendation was **most recently** made?
   _____ years old

9. Has your physician **ever** recommended you have a test for colon cancer because you had **symptoms**? Symptoms might include: abdominal pain, change in bowel habits, blood in the stool, decreased appetite, or low blood cell count (anemia). ☐$_1$ Yes ☐$_0$ No (go to question 10) ☐$_9$ Unsure

   9a. Approximately how old were you when this recommendation was **first** made?
   _____ years old

   9b. Approximately how old were you when this recommendation was **most recently** made?
   _____ years old

Fecal Occult Cards

10. Has your doctor ever **recommended** the **take-home** three-card fecal occult blood test? (cards are taken home, samples from 3 successive stools are placed on cards; cards are returned to the doctor's office or lab and checked for blood)

    ☐$_1$ Yes (Go to question 11)

    ☐$_0$ No
    ☐$_9$ Unsure } (Go to question 12)

11. Did you **return** the fecal occult cards? ☐$_1$ Yes ☐$_0$ No ☐$_9$ Unsure

    11a. Approximately when did you return the cards?

    ☐$_1$ Within the past year (Ineligible)

    ☐$_0$ More than 1 year ago

Fecal Immunochemical Test

12. Has your doctor ever **recommended** the **take-home** fecal immunochemical blood test? (taken home, a sample from a single stool is placed in the container provided, the kit is returned to the doctor's office or lab and checked for blood.)

    ☐$_1$ Yes (Go to question 13)

    ☐$_0$ No

    ☐$_9$ Unsure (Go to question 14)

13. Did you **return** the fecal immunochemical test kit? ☐$_1$ Yes ☐$_0$ No ☐$_9$ Unsure

    13a. Approximately when did you return the cards?

    ☐$_1$ Within the past year (Ineligible)

    ☐$_0$ More than 1 year ago

Flexible Sigmoidoscopy

14. Has your doctor ever **recommended** a flexible sigmoidoscopy? (where you have an enema to clear the last part of the colon and a flexible lighted tube is inserted into the rectum and 1/3 of the colon is examined; patients are not sedated)

☐$_1$ Yes (Go to question 15)

☐$_0$ No

☐$_9$ Unsure (Go to question 16)

15. Did you have the flexible sigmoidoscopy? ☐$_1$ Yes ☐$_0$ No (Go to question 16) ☐$_9$ Unsure

15a. Approximately how long ago was the flexible sigmoidoscopy?

☐$_1$ Within the past 5 years (Ineligible)

☐$_0$ More than 5 years ago

15b. Why did you have the flexible sigmoidoscopy? (Check all that apply)

a) ☐ Screening test/part of physical exam
b) ☐ My doctor recommended it
c) ☐ I don't remember
d) ☐ Change in bowel habits
e) ☐ Blood in my stool or positive fecal occult blood test
f) ☐ Low red blood cell count (anemia)
g) ☐ Abdominal pain
h) ☐ Decreased appetite
i) ☐ Weight loss
j) ☐ Family history of colon cancer
k) ☐ Personal history of colon
l) ☐ cancer Personal history of colon polyp
m) ☐ Personal history of inflammatory bowel disease

Colonoscopy

16. Has your doctor ever **recommended** a colonoscopy? (where you drink a solution the day before, to help clear your bowels, are put to sleep for the procedure, and a flexible tube examines the entire colon; someone needs to drive you home afterwards)

☐$_1$ Yes (Go to question 17)

☐$_0$ No
☐$_9$ Unsure } (Go to question 18)

17. Did you have a colonoscopy? ☐$_1$ Yes ☐$_0$ No (Go to question 18)
☐$_9$ Unsure

17a. Approximately how long ago was this? ☐$_1$ Within the past 10 years (Ineligible)
☐$_0$ More than 10 years ago

17b. Why did you have the colonoscopy? (Check all that apply)

a) ☐ Screening test/part of physical exam
b) ☐ My doctor recommended it
c) ☐ I don't remember
d) ☐ Change in bowel habits
e) ☐ Blood in my stool or positive fecal occult blood test
f) ☐ Low red blood cell count (anemia)
g) ☐ Abdominal pain
h) ☐ Decreased appetite
i) ☐ Weight loss
j) ☐ Family history of colon cancer
k) ☐ Personal history of colon
l) ☐ cancer Personal history of colon polyp
m) ☐ Personal history of inflammatory bowel disease

Barium Enema

18. Has your doctor **recommended** a barium enema? (where the colon is filled with barium and an x-ray is taken)

☐$_1$ Yes (Go to question 19)

☐$_0$ No
☐$_9$ Unsure } (Go to question 20)

19. Did you have a barium enema? ☐$_1$ Yes ☐$_0$ No (Go to question 20)
☐$_9$ Unsure

19a. Approximately how long ago was this? ☐$_1$ Within the past 5 years (Ineligible)
☐$_0$ More than 5 years ago

19b. Why did you have the barium enema? (Check all that apply)

a) ☐ Screening test/part of physical exam
b) ☐ My doctor recommended it
c) ☐ I don't remember
d) ☐ Change in bowel habits
e) ☐ Blood in my stool or positive fecal occult blood test
f) ☐ Low red blood cell count (anemia)
g) ☐ Abdominal pain
h) ☐ Decreased appetite
i) ☐ Weight loss
j) ☐ Family history of colon cancer
k) ☐ Personal history of colon
l) ☐ cancer Personal history of colon polyp
m) ☐ Personal history of inflammatory bowel disease

Demographics

Please answer each question below by checking the box or filling the blank with your best answer.

1. What is your marital status? **(check one box)**
   - ☐1 Single
   - ☐2 Married/marriage-like relationship
   - ☐3 Widowed

2. Do you consider yourself Hispanic / Latino?
   - ☐1 Yes
   - ☐2 No
   - ☐3 Don't know

3. What is your race?
   - ☐1 White
   - ☐2 Black
   - ☐3 Asian
   - ☐4 Pacific Islander
   - ☐5 Asian Indian
   - ☐6 Other _____

4. How much schooling have you had? **(check one)**
   - ☐1 8 grades or less
   - ☐2 Some high school
   - ☐3 High school graduate or GED
   - ☐4 Some college or technical school
   - ☐5 College graduate(bachelor's degree)
   - ☐6 Graduate degree

5. Do you have a DVD player at home?
   - ☐1 Yes ☐2 No

6. How would you describe the insurance plan(s) you have had **in the past 12 months? (check all that apply)**
   - ☐1 An individual plan – the member pays for the plan premium
   - ☐2 A group plan through an employer, union, etc. – the employer pays all or part of the plan premium
   - ☐3 U.S. Governmental Health Plan (e.g., Military, CHAMPUS, VA)
   - ☐4 Medicaid
   - ☐5 Medicare
   - ☐6 Medicare plus supplement I pay for
   - ☐7 I have not had an insurance plan in the past 12 months

7. Which of the categories best describes your total annual **combined** household income from **all** sources? **(check one box)**.
   - ☐1 Less than $20,000
   - ☐2 $20,000 to $39,999
   - ☐3 $40,000 to $59,999
   - ☐4 $60,000 to $79,999
   - ☐5 $80,000 to $99,999
   - ☐6 Greater than $100,000

THANK YOU FOR PARTICIPATING IN THIS STUDY!

REFERENCES

American Cancer Society. (2014). *Cancer facts & figures 2014*. Atlanta: American Cancer Society.

Anderson, A., Henry, K., Samadder, N. J., Merrill, R., & Kinney, A. (2013). Rural vs urban residence affects risk-appropriate colorectal cancer screening. *Clinical Gastroenterology and Hepatology, 11*(5), 526-533.

Baier, M., Calonge, N., Cutter, G., McClatchey, M., Schoentgen, S., Hines, S., et al. (2000). Validity of self-reported colorectal cancer screening behavior. *Cancer Epidemiology, Biomarkers Prevention, 9*(2), 229-32.

Bastani, R., Glenn, B., Maxwell, A., Ganz, P., Mojica, C., & Chang, L. C. (2008). Validation of self-reported colorectal cancer (CRC) screening in a study of ethnically diverse first-degree relatives of CRC cases. *Cancer Epidemiology, Biomarkers Prevention, 17*(4), 791-8.

Beebe, T., Jenkins, S., Anderson, K., Davern, M., & Rockwood, T. (2008). The effects of survey mode and asking about future intentions on self-reports of colorectal cancer screening. *Cancer Epidemiology, Biomarkers Prevention, 17*(4), 785-790.

Beebe, T., Ziegenfuss, J., Jenkins, S., Lackore, K., & Johnson, T. (2014). Survey mode and asking about future intentions did not impact self-reported colorectal cancer screening accuracy. *BMC Medical Research Methodology, 14*, 19-25.

Beydoun, H., & Beydoun, M. (2008). Predictors of colorectal cancer screening behaviors among average-risk older adults in the United States. *Cancer Causes & Control, 19*(4), 339-359.

Centers for Disease Control and Prevention. (2014). *Behavioral risk factor surveillance system.* Retrieved March 30, 2014, from http://www.cdc.gov/brfss/.

Centers for Disease Control and Prevention. (2014). *National health interview survey.* Retrieved March 30, 2014, from http://www.cdc.gov/nchs/nhis.htm.

Coughlin, S., & Thompson, T. (2004). Colorectal cancer screening practices among men and women in rural and nonrural areas of the United States, 1999. *The Journal of Rural Health, 20*(2), 118-124.

Daly, J., Xu, Y., Ely, J., & Levy, B. (2012). A randomized colorectal cancer screening intervention trial in the Iowa Research Network: Study recruitment methods and baseline results. *J Am Board Fam Med*, *25*(1), 63-72.

Dillman, D. (2000). *Mail and internet surveys: The tailored design method*. New York: Wiley.

Ferrante, J., Ohman Strickland, P., Hahn, K., Hudson, S., Shaw, E., Crosson, J., et al. (2008). Self-report versus medical records for assessing cancer-preventive services delivery. *Cancer Epidemiology, Biomarkers Prevention, 17*(11), 2987-2994.

Fiscella, K., Holt, K., Meldrum, S., & Franks, P. (2006). Disparities in preventive procedures: Comparisons of self-report and Medicare claims data. *BMC Health Services Research, 6*, 122-129.

Griffin, J., Burgess, D., Vernon, S., Friedemann Sanchez, G., Powell, A., van Ryn, M., et al. (2009). Are gender differences in colorectal cancer screening rates due to differences in self-reporting? *Preventive Medicine, 49*(5), 436-441.

Hall, H. I., Van Den Eeden, S. K., Tolsma, D., Rardin, K., Thompson, T., Hughes Sinclair, A., et al. (2004). Testing for prostate and colorectal cancer: Comparison of self-report and medical record audit. *Preventive Medicine, 39*(1), 27-35.

Hart, L. G., Larson, E., & Lishner, D. (2005). Rural definitions for health policy and research. *American Journal of Public Health, 95*(7), 1149-1155 (https://depts.washington.edu/uwruca/index.php).

Hoffmeister, M., Chang Claude, J., & Brenner, H. (2007). Validity of self-reported endoscopies of the large bowel and implications for estimates of colorectal cancer risk. *American Journal of Epidemiology, 166*(2), 130-136.

Hosmer, D.W., & Lemeshow, S. *Applied Logistic Regression*, 2nd edition, New York: John Wiley & Sons, Inc.

Jones CA, Parker TS, Ahearn M, Mishra AK, Variyam JN. (2009). *Health status and health care access of farm and rural populations.* No. Economic Information Bulletin No. 57) U.S. Department of Agriculture, Economic Research Service.

Jones, R., Mongin, S., Lazovich, D., Church, T., & Yeazel, M. (2008). Validity of four self-reported colorectal cancer screening modalities in a general population: Differences over time and by intervention assignment. *Cancer Epidemiology, Biomarkers Prevention, 17*(4), 777-84.

Khoja, S., McGregor, S. E., & Hilsden, R. (2007). Validation of self-reported history of colorectal cancer screening. *Canadian Family Physician, 53*(7), 1192-7.

Levin, B., Lieberman, D., McFarland, B., Andrews, K., Brooks, D., Bond, J., et al. (2008). Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology, 134*(5), 1570-95.

Levin, B., Lieberman, D., McFarland, B., Smith, R, Brooks, D., Andrews, K., et al. (2008). Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Cancer, 58*(3), 130-60.

Levy, B. T., Daly, J., Xu, Y., & Ely, J. (2012). Mailed fecal immunochemical tests plus educational materials to improve colon cancer screening rates in Iowa Research Network (IRENE) practices. *American Board of Family Medicine, 25*(1), 73-82.

Levy, B. T., Dawson, J., Hartz, A. J., & James, P. A. (2006). Colorectal cancer testing among patients cared for by Iowa family physicians. *American Journal of Preventive Medicine, 31*(3), 193-201.

Lipkus, I. M., Lyna, P. R., & Rimer, B. K. (2000). Colorectal cancer risk perceptions and screening intentions in a minority population. *Journal of the National Medical Association, 92*(10), 492-500.

Lipkus, I., Samsa, G., Dement, J., Skinner, C., Green, L. S. G., Pompeii, L., et al. (2003). Accuracy of self-reports of fecal occult blood tests and test results among individuals in the carpentry trade. *Preventive Medicine, 37*(5), 513-519.

Lynch, B., Youlden, D., Fritschi, L., Newman, B., Pakenham, K., Leggett, B., et al. (2008). Self-reported information on the diagnosis of colorectal cancer was reliable but not necessarily valid. *Journal of Clinical Epidemiology, 61*(5), 498-504.

Madlensky, L., McLaughlin, J., & Goel, V. (2003). A comparison of self-reported colorectal cancer screening with medical records. *Cancer Epidemiology, Biomarkers Prevention, 12*(7), 656-9.

McLafferty, S., & Wang, F. (2009). Rural reversal? rural-urban disparities in late-stage cancer risk in Illinois. *Cancer, 115*(12), 2755-2764.

Meissner, H., Breen, N., Klabunde, C., & Vernon, S. (2006). Patterns of colorectal cancer screening uptake among men and women in the united states. Cancer Epidemiology, Biomarkers & Prevention, 15(2), 389-394.

Meissner, H., Vernon, S., Rimer, B., Wilson, K., Rakowski, W., Briss, P., et al. (2004). The future of research that promotes cancer screening. *Cancer, 101*(5 Suppl), 1251-9.

Mosen, D., Feldstein, A., Perrin, N., Rosales, A. G., Smith, D., Liles, E., et al. (2013). More comprehensive discussion of CRC screening associated with higher screening. *The American Journal of Managed Care, 19*(4), 265-271.

National Cancer Institute. (2014). *Health information national trends survey*. Retrieved April 5, 2014, from http://hints.cancer.gov/.

Pagano, M., $ Gauvreau. (2000). *Principles of Biostatistics*, 2nd edition, Pacific Grove, CA: Duxbury.

Partin, M., Grill, J., Noorbaloochi, S., Powell, A., Burgess, D., Vernon, S., et al. (2008). Validation of self-reported colorectal cancer screening behavior from a mixed-mode survey of veterans. *Cancer Epidemiology, Biomarkers Prevention, 17*(4), 768-76.

Peek-Asa, C., Wallis, A., Harland, K., Beyer, K., Dickey, P., & Saftlas, A. (2011). Rural disparity in domestic violence prevalence and access to resources. *Journal of Women's Health, 20*(11), 1743-1749.

Powe, B., & Cooper, D. (2008). Self-reported cancer screening rates versus medical record documentation: Incongruence, specificity, and sensitivity for African American women. *Oncology Nursing Forum, 35*(2), 199-204.

Rauscher, G., Johnson, T., Cho, Y., & Walk, J. (2008). Accuracy of self-reported cancer-screening histories: A meta-analysis. *Cancer Epidemiology, Biomarkers Prevention, 17*(4), 748-757.

Reiter, P., Katz, M., Oliveri, J., Young, G., Llanos, A., & Paskett, E. (2013). Validation of self-reported colorectal cancer screening behaviors among Appalachian residents. *Public Health Nursing, 30*(4), 312-322.

Rim, S., Joseph, D., Steele, C. B., Thompson, T., & Seeff, L. (2011). Colorectal cancer screening - United States, 2002, 2004, 2006, and 2008. MMWR. Surveillance Summaries, 60 Suppl, 42-46.

Schenck, A., Klabunde, C., Warren, J., Peacock, S., Davis, W., Hawley, S., et al. (2008). Evaluation of claims, medical records, and self-report for measuring fecal occult blood testing among Medicare enrollees in fee for service. *Cancer Epidemiology, Biomarkers Prevention, 17*(4), 799-804.

Schoen, R., Weissfeld, J., Trauth, J., Ling, B., & Hayran, M. (2002). A population-based, community estimate of total colon examination: The impact on compliance with screening for colorectal cancer. *The American Journal of Gastroenterology, 97*(2), 446-451.

Shokar, N., Vernon, S., & Carlson, C. (2011). Validity of self-reported colorectal cancer test use in different racial/ethnic groups. *Family Practice, 28*, 683-688.

State Health Registry of Iowa. *Cancer in Iowa: 2014*. Iowa City: University of Iowa College of Public Health (http://cph.uiowa.edu/shri/pubs/pdf/Cancer_2014.pdf).

Sudman, S., Warnecke, R., Johnson, T., & O'Rourke, D. (1994). *Cognitive aspects of reporting cancer prevention examinations and tests.* No. Vital Health Stat 6(7)). Washington, D.C.: National Center for Health Statistics, DHHS publication; no. (PHS) 94-1082.

Thompson, N, Lynch, C, West, M et al. (2006). *Increasing colorectal cancer screening in Iowa: Needs and strategies for improvement*. Iowa City, IA: Iowa Department of Public Health and the University of Iowa College of Public Health (http://www.canceriowa.org/Files/News/crcmonograph.aspx).

Tisnado, D., Adams, J., Liu, H., Damberg, C., Chen, W., Hu, F., et al. (2006). What is the concordance between the medical record and patient self-report as data sources for ambulatory care? *Medical Care, 44*(2), 132-140.

U.S. Preventive Services Task Force (2008). Screening for colorectal cancer: U.S. preventive services task force recommendation statement. *Annals of Internal Medicine, 149*(9), 627-637.

U.S. Bureau of the Census. *U.S. census bureau delivers Iowa's 2010 census population totals, including first look at race and Hispanic origin data for legislative redistricting*. Retrieved April 27, 2014 from http://www.census.gov/newsroom/releases/archives/2010_census/cb11-cn24.html.

Vernon, S., Abotchie, P., McQueen, A., White, A., Eberth, J., & Coan, S. (2012). Is the accuracy of self-reported colorectal cancer screening associated with social desirability? *Cancer Epidemiology, Biomarkers Prevention, 21*(1), 61-65.

Vernon, S., Briss, P., Tiro, J., & Warnecke, R. (2004). Some methodologic lessons learned from cancer screening research. *Cancer, 101*(5 Suppl), 1131-45.

Vernon, S., Meissner, H., Klabunde, C., Rimer, B., Ahnen, D., Bastani, R., et al. (2004). Measures for ascertaining use of colorectal cancer screening in behavioral, health services, and epidemiologic research. *Cancer Epidemiology, Biomarkers Prevention, 13*(6), 898-905.

Vernon, S., Tiro, J., Vojvodic, R., Coan, S., Diamond, P., Greisinger, A., et al. (2008). Reliability and validity of a questionnaire to measure colorectal cancer screening behaviors: Does mode of survey administration matter? *Cancer Epidemiology, Biomarkers Prevention, 17*(4), 758-67.

WWAMI Rural Health Research Center. *Using RUCA data*. Retrieved April 27, 2014 from http://www.depts.washington.edu/uwruca/ruca-uses.php.

White, A., Vernon, S., Eberth, J., Tiro, J, Coan, S., Abotchie, P., et al. (2013). Correlates of self-reported colorectal cancer screening accuracy in a multi-specialty medical group practice. *Open J Epidemiology, 3*(1), 20-24.

Zapka, J. (2008). Innovative provider- and health system-directed approaches to improving colorectal cancer screening delivery. *Medical Care, 46*(9 Suppl 1), S62-S67.

Zauber, A., Winawer, S., O'Brien, M., Lansdorp Vogelaar, I., van Ballegooijen, M., Hankey, B., et al. (2012). Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *The New England Journal of Medicine, 366*(8), 687-696.