Measures Used in Studies of Informed Decision Making About Cancer Screening: A Systematic Review

Patricia Dolan Mullen, Dr.P.H.
University of Texas School of Public Health

Jennifer Dacey Allen, R.N., M.P.H., Sc.D.
Dana-Farber Cancer Institute and Harvard School of Public Health

Karen Glanz, Ph.D., M.P.H.
Emory University, Rollins School of Public Health

Maria E. Fernandez, Ph.D.
University of Texas School of Public Health

Deborah J. Bowen, Ph.D.
University of Washington, School of Public Health and Community Medicine and Fred Hutchinson Cancer Research Center

Sandi L. Pruitt, M.P.H.
University of Texas School of Public Health

Beth A. Glenn, Ph.D.
University of California at Los Angeles, School of Public Health

Michael Pignone, M.D., M.P.H
University of North Carolina at Chapel Hill, School of Medicine

ABSTRACT

Interventions to promote informed decision making (IDM) for cancer screening are increasingly common. The resulting body of literature provides an opportunity for a systematic review of measures in use. We searched standard databases for intervention trials and other studies of screening decisions and decision aids, finding 2,110 unique citations (most with abstracts) that we reduced to 104 full-text articles; 36 studies met inclusion criteria (prostate = 20, colorectal = 9, breast = 6, cervical = 1). Two independent coders abstracted data on study characteristics, constructs, and measures. Our findings revealed that most studies measured screening (or intention) and knowledge; fewer measured recommended IDM-related constructs and none measured all outcomes proposed for evaluating IDM interventions. Validity and reliability of measures received inadequate attention in study reports, and conceptual overlap exists among measures. Few IDM measures have been developed/carefully adapted from treatment measures and tested for cancer screening or in diverse populations. We recommend that new and in-progress studies emphasize outcomes beyond knowledge—participation in decision making according to personal preference, satisfaction with the process, and consistency between decisions and values. Also needed is better use of theory to guide conceptualization and operationalization of measures, greater attention to reliability and validity (particularly in diverse populations), more thorough reporting of sources and operating characteristics of measures, and increased emphasis and resources focused on these issues by funders, researchers, and journal editors.

(**Ann Behav Med** 2006, *32*(3):188–201)

INTRODUCTION

There are growing efforts to use informed decision-making processes to help patients make decisions about cancer screening, as well as increased desire on the part of patients to be involved

Research for this publication was supported by the Centers for Disease Control and Prevention (CDC) and the National Cancer Institute (NCI) cooperative agreements for the Cancer Prevention and Control Research Network (CPCRN) at Emory University, Rollins School of Public Health (1-U48 DP00043); Harvard University/Boston University Schools of Public Health (1-U48-DP-000064); University of California at Los Angeles, School of Public Health (1-U48-DP-000056); University of North Carolina at Chapel Hill (UNC), Center for Health Promotion and Disease Prevention (1-U48-DP-000059); University of Texas School of Public Health (UTSPH) (1-U48-DP-000057); and University of Washington, School of Public Health and Community Medicine (1-U48-DP-000050). Ms. Pruitt was supported by NCI training grant R25CA57713 (UTSPH).

We thank Kathleen A. McGraw, M.A., M.L.S., for conducting the literature search, Rebecca Williams, Ph.D., for managing the citation database, and Sarah Walden for administrative assistance (CPCRN Coordinating Center, UNC); Sandra Tyson, M.A., and Belinda Flores (UTSPH CPCRN), Michelle Davis, and Suzanne Kneuper, M.A. (CDC special interest project 23–04 to UTSPH CPCRN) for assistance in checking coding and tables and locating background reports on measures; and Katherine Regan Sterba, Ph.D., NCI-UTSPH Postdoctoral Fellow (R25CA57713) for her comments on the article.

Reprint Address: P. D. Mullen, Dr.PH., Center for Health Promotion and Prevention Research, University of Texas, School of Public Health, 7000 Fannin, UCT Suite 2522, Houston, TX 77030. E-mail: Patricia.D.Mullen@uth.tmc.edu

in decision making (1,2). Screening tests for breast, cervical, colorectal, and prostate cancer are widely used and, depending on available test options and the strength of the evidence at a particular point in time, can present differing types of decisions that could benefit from patient IDM. IDM is generally defined as the process that patients go through to make a decision about engaging in a medical or health-related procedure or activity considering benefits, harms, risks, health improvements, the match between these properties and personal values and preferences, and understanding the uncertainty and limitations of the procedures. Patients then often undergo a process of shared decision making with a provider to make a final decision (3).

Decision aids are designed to help people make specific and deliberative choices among options by providing (at a minimum) information on the options and outcomes relevant to their health status (4,5). Recent reviews (3–6) have documented the effects of these interventions on multiple potential intermediate and more distal outcomes of decision-making processes for treatment and screening. In their review for the U.S. Community Preventive Services Task Force, Briss and associates concluded that the evidence for using decision aids to improve patients' decision processes in cancer screening decisions was insufficient to recommend current use (3). Specific recommendations included standardization of outcome measures for decision-making studies and further research and conceptualization of the decision-making process in diverse groups.

We present here the results of a systematic review of measures used to evaluate decision tools and to understand screening decisions. For this review, we drew on the definition of IDM adopted by the U.S. Community Preventive Services Task Force (3), in collaboration with a workgroup from the U.S. [Clinical] Preventive Services Task Force, which focused on shared decision making, as a subset of IDM (7). This work benefited from previous work by investigators and theorists who have done much to define the field (4,8–11). The definition and constructs suggested in the analytic framework proposed by Briss and associates (3) guided our selection of measures to evaluate.

Having relevant, high-quality measures for the outcomes and processes of IDM is a critical step to improving our understanding of the findings of new cancer screening studies. Pioneering work by the investigators who were represented in reports identified in the course of our literature search provides an opportunity to assess the state of measures currently used and to consider next steps to strengthen them. Our aims for this review, therefore, are to (a) describe the measures used in studies of IDM for cancer screening, including available information about their characteristics and performance; (b) determine the extent to which the constructs that have been measured represent the definition of IDM adopted in the U.S. Community Preventive Services Task Force review; and (c) draw on the expertise of five Centers for Disease Control and Prevention and the National Cancer Institute–funded Cancer Prevention and Control Research Networks represented by the authors to provide recommendations to improve the measurement of IDM constructs in the next generation of studies.

METHOD

Relevant Constructs

The definition of IDM in the U.S. Community Preventive Services Task Force (The Task Force) review provided an a priori organization of constructs to focus our review (Table 1). Measures found in the studies included in this review were categorized by these constructs using the language of the authors of the primary studies and judgment by the authors of this review.

Eligibility Criteria for Primary Studies

The Task Force review's eligibility criteria were a reference point for our eligibility criteria. These were (a) English language, (b) primary study rather than a guideline or review, (c) took place in a developed country, (d) provided information on one or more outcomes related to the analytic framework, (e) met the Task Force review team's definition of the intervention, and (f) compared an exposed group with a group not exposed (comparisons could be in concurrent or in the same group over a period of time (3). We adopted Criteria a to d and included all 13 intervention trials from the Task Force review. Because we wanted to capture information about constructs and measures of proximal and distal outcomes of IDM and not the efficacy or effectiveness of IDM interventions, we broadened Criteria e and f to include one group pretest–posttest trials, baseline measures from intervention trials (when trial results were not available), other cross-sectional surveys, and qualitative studies for intervention development. Finally, eligible studies were focused on *decision making* for first-line cancer screening (prostate, colorectal, breast, or cervical cancer), including testing for human papillomavirus (HPV). Based on the results of the earlier literature search for the Task Force review, we anticipated that the focus on decision making would reduce the likelihood of capturing studies of breast, cervical, and colorectal cancer screening interventions or studies of screening predictors—studies in which persuasion rather than decision making was the chief objective.

Search Strategies

The literature search was conducted on three electronic databases (PubMed, CINAHL, and PsycINFO) by a health sciences research librarian in October 2005. The search was an adaptation and expansion of the search strategy used in the Task Force review (3) (search terms from Peter Briss, personal communication to Mullen, October 17, 2005). PubMed (NCBI) was searched using Medical Subject Headings (MeSH) and text words with three sets of search terms (Cancer/Examinations, Decision Making, and Screening). No publication date limits were set; PubMed includes citations dating back to 1950. The PubMed search was then adapted for CINAHL (EBSCO), dating back to 1982, and PsycINFO (WebSPIRS), dating back to 1887, using each database's unique thesaurus of subject headings. (The complete search strategy is available on http://www. cpcrn.org.) The number of citations (k) from PubMed ($k = 1785$), CINAHL ($k = 571$), and PsycINFO ($k = 280$), respectively, were downloaded to a bibliographic software database from which duplicate citations ($k =$

TABLE 1
Citations by IDM Definition, Construct, Cancer Site, and Study Type

Definition: IDM Occurs When an Individual …	*Construct*	*Cancer Site*		
		Prostate[a]	*Colorectal*[b]	*Breast*[c,d]
Understands the test, the condition, personal risks, uncertainties	Knowledge	Trials: (12–24) Other: (25)	Trials: (26–30)	Trials: (31,32) Other:(33)
	Perceived threat, worry	Trials: (23,34)	Other: (23)	—
	Perceived risk, susceptibility	Trials: (18,19,35) Other: (36)	Other: (37)	Trial: (32) Other: (38)
	Perceived severity/ seriousness	Trials: (18,35) Other: (36)	—	Other: (38)
Considers preferences	Decisional balance	Trials: (18,19)	Other: (37)	Trials: (31,32) Other: (33)
	Perceived benefits	Trials: (19,35) Other: (36)	Trial: (30) Others: (37,39)	Trial: (32) Others: (38,40)
	Perceived barriers	Trial: (35)	Others: (37,39)	Others: (33)
	Attitude re: the test/screening	—	Other: (37)	Other: (40)
	Role preference	Trials: (16–18,34) Others: (19,41)	Trial: (29) Others: (37,39)	Others: (33,38,40,42)
	Utilities, values	Trial: (35) Others: (43,44)	—	—
	Treatment preference (if diagnosed)	Trials: (12,13,17,22)	—	—
Participates in decision at a personally desirable level	Decisional self-efficacy	Trials: (16–19) Other: (36)	Trial: (27)	—
	Discussion with clinician (intention)	Trial: (15)	—	Other: (33)
	Discussion with clinician, role performance	Trials: (15,20,22,34)	Trials: (29,45) Other: (37)	—
Makes a decision consistent with values	Test preference	—	Trials: (30,46)	—
	Screening (intention)	Trials: (12–14,17,18, 20,21,23,24,35,41)	Trials: (26,27,30,45)	Other: (33)
	Screening	Trials: (12,13,16,17,20–22,34) Others: (25,36)	Trials: (26–29,45) Other: (37)	Trial: (32) Others: (40,47)
	Satisfaction with the decision	Trial: (21)	—	—
	Decisional conflict	Trials: (18,19,34)	Trial: (29)	Other: (40)

Note. IDM = informed decision making; Others = cross-sectional surveys, the prospective cohort study, qualitative studies.
[a]Trials, $k = 16$; Others, $k = 5$. [b]Trials, $k = 7$; Others, $k = 2$. [c]Trials, $k = 2$; Others, $k = 4 + 1$. [d]The cervical cancer study has been included in this column: (38).

340) and articles clearly off topic ($k = 286$), such as those about genetic screening, were removed.

Review of Citations

Review and selection of potentially relevant articles used a stepwise process. First, the list of potentially eligible citations and abstracts was divided among the authors who independently selected articles for full-text review. Selection by any author led to retrieval of the full-text article. Second, every full-text article was reviewed independently by two authors with discrepancies resolved by discussion and consensus of the whole group. Last, the initial list of potentially eligible citations and the final set of eligible articles was compared with the Task Force review (3) to ensure that our search captured relevant articles.

Coding

Selected studies were read and coded by one of the authors, using a uniform coding form, and read by a second author who independently checked the code sheet. Discrepancies were resolved by discussion and consensus. Studies were coded for standard characteristics including study design, intervention, sample, setting, screening test, constructs measured, specific information about the measures, their role in analyses, and indications of their reliability and validity. In keeping with our aims,

the primary emphasis was on psychosocial and behavioral constructs relevant to IDM (Table 1).

Methods of Analysis

Measure-specific information from the primary studies was entered into evidence tables for each relevant construct (tables available from Mullen). For the knowledge construct, an additional table was created to display domains by cancer site (table available from Mullen). Counts, examples, and general descriptions were then generated from the tables. Standards of comparison used to interpret the results in the Discussion section included the Task Force review's logic model, specifically cited definitions and commentaries from experts on decision making, and generally accepted measurement practice (48–50).

RESULTS

Studies Identified

A comparison of the 2,010 potentially relevant citations from our search with citations from the Task Force review confirmed that the new search had found all of the target studies and citations except one—a breast cancer screening decision study that did not have any variation of the term *screen(ing)* in the title or abstract and was not indexed to the MeSH term *mass screening* (32). We added this citation so that all studies identified by Briss and associates (3) were included in our database. After the initial review of 1,709 citations with abstracts and 301 citations with titles only, we retained 104 (5.2%). Review of the full texts of these citations identified 37 studies that met our inclusion criteria. One study reporting an analysis from the Health Information National Trends Survey (51) was excluded because no IDM construct other than screening was reported. Thus, the final number of studies was 36. To avoid confusion in the case of a single article that contained reports of two different studies, we assigned each its own citation (12,13). Primary citations for the included studies appear in Table 1. Secondary citations, for example, the report of an early follow-up, are listed on the evidence table describing study design, sample, interventions, and other study characteristics available on http://www.cpcrn.org (32,52–56).

More than half of the studies (k = 20/36) were about prostate cancer screening, the cancer site with inconclusive evidence for screening and where the decision about *whether* to be screened has been the main focus of IDM research in cancer screening (Table 1). The next largest group, colorectal cancer (CRC) screening (k = 9/36), has the requisite evidence of the impact of screening on mortality but presents many patients with the decision of which test to have. Although screening mammography for women ages 40 to 50 at normal risk for breast cancer has been viewed by some expert groups as appropriate for IDM, few breast and cervical cancer screening studies have focused on IDM or decision making and thus, as expected, few of these studies were identified by our search terms.

The majority of the studies (k = 25/36) were intervention trials (Table 1). Prostate cancer studies typically focused on the prostate-specific antigen (PSA) test; six also incorporated digital rectal examination. Of the CRC studies, most focused on fecal occult blood testing (FOBT), although three included flexible sigmoidoscopy, and one of those also included colonoscopy. No study included more than one cancer site. Most were conducted in the United States; 6 were located in Canada or Australia. Publication dates ranged from 1993 through 2005, with 14 studies published in 2004 or 2005. All were articles in peer-reviewed journals except three dissertations that had not been published by the time of our search (26,38,47). More than half of the intervention trials used random assignment of individuals or clusters to intervention and control groups. Eight other studies were cross-sectional surveys in convenience samples; 2 others used semistructured interviews and qualitative analysis methods.

Overview of Constructs Measured

We identified and coded 150 measures of relevant IDM constructs found in primary and secondary citations of the included studies. Study citations for each construct by cancer site and study type are shown in Table 1. (In cases of multiple measures of the same construct, the citation appears only once.) Some measures were omitted because they could not be matched with a single construct—two highly heterogeneous belief scales (24,27), the item, "all men should be screened" (22), and what information doctors should tell women about breast cancer screening (40). We also excluded health locus of control (19), urologic symptoms (19,20,34), quality of life (41), life satisfaction (41), health information seeking (17,21,42), numeracy (31), general satisfaction with doctor–patient communication (28,40) and self-care practices (26), intention to talk to the family about their being screened, and self-efficacy regarding this discussion (27). After these exclusions, 131 measures remained.

Across cancer sites, the constructs measured most frequently were screening, including intention regarding screening (k = 30/36) and knowledge (k = 27/36), the most distal and most proximal measures, respectively, in the conceptual framework (Table 1). Next in order of frequency of measurement were (a) role preference (k = 11); (b) cancer threat, including perceived risk and/or severity (k = 10); (c) discussion of screening with the clinician, including intention to discuss screening with the clinician (k = 8); (d) decisional balance, including benefits or barriers or both (k = 6); (e) self-efficacy (k = 6); and (f) decisional conflict (k = 5). Few studies measured other decision-making process or outcome constructs, including values/utilities (k = 3), role performance consistent with role preference (k = 1), or satisfaction with the decision (k = 1) or decision-making process (k = 0).

Several differences across cancer sites in the relative frequency of constructs measured reflect differences in the type of decision being made. Only prostate cancer screening studies measured treatment preference, and values/utilities, and decision self-efficacy (i.e., confidence in ability to make an informed decision; Table 1). CRC studies were expected to include a measure of test preference (among test options), and both of the CRC trials that included multiple tests did. No other patterns were noted (Table 1).

The source of measures in the included studies was often unclear. For about half of the constructs other than screening, the description of the measure was either not accompanied by a

citation (frequently) or the measure was described as created by the authors (infrequently). For the other half, the description or mention of the measure was accompanied by a citation, often without an indication or description of adaptation of the measure associated with the citation. The range of adaptation and specificity of the descriptions was very broad, as illustrated later in this article.

Knowledge

Knowledge is widely identified in the IDM literature as a critical component of IDM (12), particularly in commentaries taking the perspective of informed consent (57). Knowledge was measured in most of the trials (k = 22/25) and few of the other studies (k = 2/10) (see Table 1). The typical analytic role for knowledge in intervention trials was as an intermediate or primary outcome, as it was in 13 trials that measured knowledge at baseline and at a follow-up time soon after the intervention (14–19,21,22,24,27,31,32,41) and in 5 trials that measured knowledge at follow-up only (12,13,20,26,30). In the former studies, knowledge was available as a potential covariate to use in equating study arms; in the latter, however, the protocol had the advantage of avoiding testing effects, an important matter when the interval between measures is as brief as 2 weeks (58). In one prostate cancer trial, *self-reported knowledge* was measured at baseline and follow-up (23). One CRC trial in which knowledge was measured at baseline only reported the results as a cross-sectional survey (56) and did not measure it again at the 12-month follow-up (28).

Diversity in the knowledge domains measured in the trials appeared to have some relation to a cancer site. (Three studies reported too little information to be included in the domain review, including the study measuring self-reported knowledge [23,28,29].) Six domains were measured exclusively in the prostate cancer studies although not uniformly across prostate cancer studies (a) the natural history of the cancer (12–17,19,20,22,24), (b) treatment efficacy (12,13,16–18,20–22), (c) treatment side effects (16–19,21,25), (d) expert disagreement about mass testing (15,19,20,25), (e) physiology (14), and (f) confirmatory tests (24). Expert recommendations about the essential knowledge domains for prostate cancer screening have included a to d, f, and risk factors (57,58), suggesting the need for more attention to underrepresented domains, in particular, controversy about screening. Accuracy of screening tests was measured for prostate cancer (12,13,16–22,24), CRC (26,27), and breast cancer (10,31). Risk factors and epidemiology also were measured across cancer site (19–21,24,26,27), as was the domain, screening tests, and guidelines (12–14,24,26,27).

The number of items used in the knowledge measures ranged from a single item (the positive predictive value of FOBT) (30) to a 25-item scale (26), with a mode of 10 items. Scores for subscales and scales were typically based on the sum of correct answers, although in a few cases investigators assessed the impact of the intervention on individual items. The reported response options varied—true–false items only, with and without a "not sure" option (18,20,25,26,28,29); multiple choice only (10,12,13,16,17,21,30,31,33); 3-point agreement options (27); and mixed response formats (19,52); five studies that measured knowledge omitted information about response options (7,14,22,24,41).

Several study reports provided information about measure development, including identification of candidate items, pretesting, and refinement of scales (19,25–28,41). Methods reported for establishing content validity were asking experts what patients ought to know before being tested (19,21,25,41) and deriving the items from the content of the intervention (14). The range of domains even within cancer site, suggest content validity is in need of more attention (58).

Indicators of reliability were reported infrequently. Eight of the 22 studies with multi-item knowledge measures reported consistency reliability (Cronbach's alpha) ranging from .63 to .92. Two study reports described stability reliability (test–retest) (25,41), and one, split-half reliability (20).

Extensive developmental work and testing of a knowledge measure for use in the Prostate Cancer Screening Education Study trial (20) provides a template for developing and testing measures through steps that occurred before the trial and steps that were incorporated into collection and analysis of baseline data. Developers focused on item difficulty, item discrimination, reliability (split-half reliability), and validity (content, construct, and criterion validity) (58), and they suggested an agenda for improving measurement, including testing in more racial and ethnically diverse samples, especially African Americans, and testing across different modes of administration.

Perceived Threat

This construct, most often identified as a component of the Health Belief Model (HBM) (59,60) and the larger class of value expectancy models, was measured as a general construct in three studies, and the closely related constructs, *perceived risk* or *susceptibility* and *perceived severity*, was measured in seven additional studies (see later). None of the studies assessing cancer threat explicitly referred to a theory or model. One trial used a single item, "concern about prostate cancer" (7); another, a qualitative interview study, identified "fear of cancer," a concept appearing to reflect a combination of perceived risk and seriousness (39). The construct as measured in these two studies seemed to reflect a combination of perceived risk and seriousness of cancer. In the third study, a prostate cancer screening trial, Davison and associates (34) assessed state anxiety as an intermediate outcome of their intervention, using the 20-item State Anxiety Inventory (61), a scale with well-established psychometric properties. State anxiety is often measured in studies of decision aids for treatment (5). Although state anxiety is similar to "cancer worry," that construct has some overlap with perceived threat. Furthermore, no cancer-specific measure of worry was included in this study. The heterogeneity of these three measures provides an example of conceptual and operational inconsistencies reported in two recent reviews (62,63).

Perceived risk of having cancer, susceptibility. Perceived susceptibility of developing cancer was assessed specifically in four trials and three other studies across cancer sites (Table 1). In three studies in which perceived risk was measured explicitly

as an HBM construct (35,38) or a Protection Motivation Theory construct (36), one of which used a single item (35), all reported adapting an existing six-item HBM construct measure that had been developed for breast cancer screening (64,65). Across the studies with measures of perceived susceptibility, measures were similar in that their referent was the specific type of cancer, and risk was estimated in terms of personal risk or "people like me." The measures varied in number of items (range = 1–6), analytic role (covariate or intermediate outcome), type of risk (absolute or comparative), and time frame (next 10 years, lifetime, or not defined). Little information about reliability and validity was presented other than Cronbach's alpha for the four- and six-item measures that had explicitly drawn on the extant HBM breast cancer measures.

Rimer (66) compared the accuracy of study participants' perceived absolute 10-year risk of breast cancer with their risk as assessed by the Gail model. Gattellari and associates (18) compared perceived risk with estimated lifetime incidence of prostate cancer. The results of their comparisons, which could be viewed as an indicator of knowledge or risk perception, illustrates the conceptual overlap between knowledge and perceptions. This measure is similar to what O'Connor called "realistic expectations," the accuracy of perceived uncertainty or risk (67).

Perceived severity of cancer or seriousness of having the disease. All four studies that measured this construct also measured perceived risk (Table 1). The measures of this construct for a particular study were very similar to that study's measure of perceived risk.

Decisional Balance

Studies measuring this construct, originally developed by Janis and Mann (68), based their measures on the extensive developmental work and testing by Rakowski and associates on a measure for mammography screening (69,70), extended to the Pap testing (71). Two of the studies were of breast cancer (10,33), and two were of prostate cancer screening (18,19). The adaptation process was not described in the latter studies, although they did report data on the characteristics of the adapted measure. Measures included 2 to 16 items as cons and 3 to 10 pros, all with 5-point Likert scales (*strongly agree* to *strongly disagree*). The convention for scoring a decisional balance measure is to subtract cons from pros (based on standardized *t*-test scores) to create a net score. For breast cancer screening, a positive score is desirable, indicating that the net balance of pros and cons favors screening. For prostate cancer, however, a positive net balance would not be considered "good" or "bad" in and of itself. The issue would be its consistency with the screening decision. Other indicators that a decision was consistent with values might be more appropriate for such a comparison, for example, direct measures of utilities and values, discussed later.

Perceived benefits and barriers of screening. Four studies measured both of these constructs, and five measured benefits only (Table 1). As with perceived risk and susceptibility, two studies explicitly described benefits and barriers as HBM constructs (35,38), and a third study described them as Protection Motivation Theory. Again, these studies as well as several others that did not mention a theory or model nevertheless all cited the same source of (HMB) measures (64,65). Another study used an eight-item response efficacy measure (36). The remaining studies used either single general items (e.g., "How convinced are you of the benefits of screening?") (19) or a single topics item (e.g., "Being screened gives a feeling of control over health") (10). Benefits and barriers measures were combined in this section to suggest their potential for combination as a net balance, as they were intended to be used by HBM originators (59). The few studies that isolated a prominent issue regarding a particular test (e.g., the mortality reduction item from a CRC trial) (30) could be used across cancer screening types. The Cronbach's alpha values reported for the multi-item scales were .91 (19) and .81 (36), respectively.

General Attitude Toward a Test or Screening

Two of the studies with measures of general attitudes used single items to indicate a general leaning for or against, for example, "The benefits of FOBT outweigh the discomfort" (37) (Table 1). A second measure in the same CRC screening trial included seven items representing negative attitudes toward CRC tests. The third study took a direction that is consistent with the current emphasis on rescreening and tapped the respondent's attitude toward her most recent screening mammogram (40). Using the same stem, "For me having the screening test [a mammogram] was" with four dimensions, each with a 5-point Likert response scale, *beneficial–harmful, very important–very unimportant, very bad thing–very good thing,* and *very pleasant–very unpleasant* (Cronbach's $\alpha = .91$).

Role Preference

The terms *decision-making involvement, role preferences,* and *decision preferences* were used interchangeably in this group of studies. The essence of a widely cited definition is the extent to which the individual wants to make (or has made) a decision about a screening test by himself or herself, or to have someone else (usually the doctor) make the decision. Most studies treated this as a unidimensional ordinal construct, on a continuum ranging from active (making the decision oneself) to collaborative (with the provider) to passive (provider makes decision).

Most of the studies reviewed reported that their measures were based on Degner's Control Preferences Scale (72–74). Some investigators referred to Degner's 1997 publication about women's treatment preferences for breast cancer (73), implying that this presented evidence of reliability and validity. Degner's work has used a multistep card sort technique, however, that was used in only one of the studies in this review (41). Other studies reviewed measured role preference with a single item where respondents were asked to choose a label on a 5-point scale that best represented their desired role. Also, Degner's research had focused on the preferred role in making treatment decisions as a more stable decision-making "style." Although Frosch and associates (16,17) and several other investigators' measures also referred to "medical treatment decisions," in most of the re-

viewed studies, the referent was changed to screening for the particular cancer site of interest.

Thus, whether role preference is appropriately measured is confused by this difference between the measure-as-used in most screening studies and the construct the original measure was designed to represent. If role preference is not a relatively stable "style," then comparing a pretest role preference measure against postintervention role performance is inappropriate. On the other hand, measuring role preference and role performance at the same follow-up time could cause the respondent to adjust the answer on whichever was the second question to be more consistent with the first (75). Results from several intervention trials have indicated that desire for participation in decision making is increased after intervention (e.g., 16,34). The authors of the former study had measured role preference using a more general question with treatment as the referent, that is, tapping the "more stable" preference.

In addition, qualitative studies, and studies with a qualitative component measuring women's preference for decision making about medical testing (47,76), suggested that the "role label" approach is simplistic, and investigators should also allow for the possibility that people other than the clinician might play a role in the decision (e.g., spouse) and that preferences could differ with the decision and the factors that were relevant to the decision.

The reviewed studies reported scant information about their measure's psychometric properties. The opportunity to test discriminate or concurrent validity within the data collected for the trial or survey did not appear to have been recognized. A scenario approach to measuring role preference, scenario measures of a physician's responsibility for decision making about prostate cancer screening, was also used (19). Original sources for this approach include Deber and associates' preference-for-involvement scenarios (77) and the scenario section of the Autonomy Preference Index (78).

Utilities, Values

Utilities, or the personal importance placed on the potential benefits or risks of a given course of action, have been measured in three prostate cancer screening studies. In the intervention trial, a "risk–benefit trade-off" was assessed as a potential mediating factor. Two mechanisms have been used to measure utilities in studies of cancer screening decision making. Volk's team (43,44) employed a "time trade-off" method, which measures the importance or "utility" a respondent places on a period of perfect health, in comparison to a period of ill health. "The objective is to determine the point at which the subject is indifferent about the choice between life with a given adverse health state or a shorter period in perfect health" (43, p. 73). For example, a respondent may be asked to compare a scenario where he would live for a number of years with one of the potential short- or long-term effects of prostate cancer treatment (e.g., impotence, incontinence) versus a shorter number of years in perfect health. The questioning continues until the point of indifference is discovered. Utilities are typically scored from 0.0 (death) to 1.0 (perfect health).

A similar method was used (35) to assess a man's willingness to accept a particular probability of risk associated with treatment, given uncertain treatment benefits and potential complications. This study used one item to assess the risk–benefit trade-off. In a variant on the time trade-off method, participants were asked to rate various potential treatment complications on a continuum (or "feeling thermometer"), again with 0 representing the worst possible state (death) and 100 representing the best possible state (perfect health) (44). Inherent in both of these forms of measurement is the view that quality versus quantity of life are the values most relevant to decision making.

Treatment Preference

Four prostate cancer screening trials measured patients' treatment preferences were they to be diagnosed with prostate cancer—surgery, radiation therapy, or watchful waiting (Table 1). This particular preference plays a unique role in decisions about prostate cancer screening tests, with a preference for watchful waiting thought to be consistent with deciding not to be screened, and a preference for surgery consistent with deciding to be screened. In at least two of the studies, treatment preference watchful waiting (vs. surgery, radiation) was an intermediate outcome; no analysis was presented that tested the link between watchful waiting and not being screened (12,13).

Decisional Self-Efficacy

Measures used in the six studies reporting the decisional self-efficacy construct were either very general single items (k = 4) or 3 to 4 specific items. Two prostate cancer screening trials (18,19) measured men's confidence in their ability to make a decision about PSA testing, using a single item, "I feel I can make an informed choice about having a PSA blood test," from the Effective Decision-Making subscale of the Decisional Conflict Scale (79,80). Another pair of prostate cancer screening trials, assessed "confidence in my PSA decision," as an intermediate outcome, using a 10-point Likert scale (16,17) and also cited O'Connor (80). The two remaining studies measured other aspects of testing-related self-confidence—confidence in one's ability to complete the testing (27,36) once a decision has been made, for example, following the instructions for FOBT (e.g., overcoming personal and cultural barriers to handling feces). None of the studies used the 11-item Decisional Self-Efficacy Scale, which does include the latter aspects of decision making (79,81).

Discussion of Screening with a Clinician (Intention) and Role Performance

Patient–provider communication, typically self-reported discussion about the test or the type of cancer, was measured in seven studies, one of which also measured intention as a secondary outcome if discussion had not occurred by the time of the follow-up measures (15) (Table 1). In one of these studies, physicians were asked their opinion about the role taken by the patient, and patient–physician reports were correlated (r^2 = .47). Intention alone was measured in one other study (Table 1). With the exception of the CRC study for which the development and

fielding of baseline measures were published separately (56) from the trial results (28,56), no information about psychometric characteristics or source of the measure was reported. Katz and associates reported that the construct had been an important theme in focus groups, reduced the original five items to a three-item subscale that was used in the analysis, and had acceptable internal consistency (56).

In the studies in which the discussion measure was specified as role performance, with measures adapted from the role preference items, one trial did not compare role preference (pretest) and role performance (34). The other trial did make this comparison, calling the combined variable "role congruence" (29,34), and the third study, a survey, compared role preference and role performance ("usual" decision making about CRC screening) (37). Degner's report on breast cancer patients' role preference for treatment decisions (73) provides a model for comparing role preference and self-reported performance; she and her associates called the areas of congruence "role congruent with preferences."

Test Preference

Two studies examined patients' choice of which CRC tests to have after receiving information about their options (Table 1). The questions, developed by the authors, were not accompanied by information about any psychometric properties. Another use of test preference has been as a baseline indicator of the extent to which the patients enter the study already having made up their minds: In a hormone replacement therapy decision aid trial, the analytic role of this variable was as a moderator of outcomes—those with "polarized preferences" did not change their minds, although they were better informed (67).

Screening (Intention Regarding Screening)

Most studies in this review measured screening or intention as a proxy for screening (Table 1). Screening was usually considered a primary study outcome. In 10 intervention trials, screening was recorded or obtained from clinical records, laboratory slips, or other objective means. In the other trials, screening was assessed by self-report (10,21,28,34). Not surprisingly, the screening measure in the surveys was self-report. Eight studies used intention, stage of change, and interest in having a test or chances that they would have a test at a specific time in the future as a proxy for screening. The studies that used self-reported measures of screening reported little information about a developmental process, use of existing instruments, or psychometric information. Two studies (34,51) used measures that had been developed for another project. The self-reported screening measures were, in general, a combination of the responses of one to three branching with yes–no response formats asking if the participant had had a screening test, and if yes, the timing of that test. As for the measures of screening intention, only a few of the authors used measures that had been previously developed (18,33); one provided any psychometric information (33). In general, the intention measures consisted of a single item (on a 4- or 5-point scale), asking participants (a) how likely it is that they would have the test, (b) whether they were planning to have the test, or (c) the chances of having the test in the future, usually 12 months hence. None of the studies reviewed mentioned using a source of intention measures, such as Ajzen (82).

Satisfaction with the Decision and With the Decision-Making Process

Satisfaction with the decision, assessed in one study, was a primary outcome at the 12-month follow-up after a PSA decision aid intervention (Table 1). The measure was based on one developed to measure satisfaction with treatment decisions (83), with 6 items that assess adequacy of information, consistency of the decision with personal values, belief in ability to carry out the decision, opportunities for sufficient input into the decision, and whether the overall decision was satisfactory. Evidence for construct validity came from correlational analyses that have demonstrated statistically significant and expected relationships with the Decisional Conflict Scale and subscales (83,84). Two problems with the Satisfaction with the Decision scale for screening decisions have been recognized and may have reduced enthusiasm for its use. The first is the relevance of this construct to an easily reversed decision, in contrast to treatment decisions, a point that has been discussed among researchers conducting studies of shared and IDM (R. Volk, personal communication to Mullen, March 26, 2006). The second problem is that measures of this construct, including three studies that used other satisfaction scales, have not registered between-group posttest differences as summarized in the Cochrane cancer decision aid review (5). Satisfaction with the decision-making process, however, has been endorsed as a relevant construct for screening decisions, and at least one 12-item scale is available (85) and was used in two studies included in the Cochrane review.

Decisional Conflict

"Decisional conflict is a state of uncertainty about the course of action taken" (83,86, p. 25) and is based conceptually in the work of Janis and Mann (68). Five studies (Table 1) measured this construct using either all or most of O'Connor's Decision Conflict Scale (86). The Decision Conflict Scale has 16 Likert-type items and three subscales: Uncertainty, Effective Decision Making, and Factors Contributing to Uncertainty. Three studies used only two subscales and omitted the Effective Decision Making subscale, apparently because they were not collecting these data after a decision was made or they were using the item as decisional self-efficacy (see earlier). The items in the scale include items that are clearly overlapping with—or perhaps more correctly, identical to—Decisional Self-Efficacy, the Satisfaction with Decision scale, and Satisfaction with the Decision Making Process. The measure has good face validity (75,87,88), well-established psychometrics—high internal consistency (0.78 to 0.92) and test–retest reliability (0.81)—and reasonable discriminant validity (86). Three of the reports in our review reported psychometric data for their studies, with similarly good reliability findings; the others referred to O'Connor's work.

DISCUSSION

These studies represent the recent considerations in measuring outcomes and other relevant constructs for studies of decision making. Our findings reveal both strengths and gaps in the constructs and measures that have been used in these studies on which we base several recommendations for strengthening future IDM studies.

Our point of reference, the Task Force Review by Briss and colleagues (3) who had drawn on the conceptualization, research, and practice of shared and IMD then available to suggest three primary outcomes for IDM studies for cancer screening (a) knowledge, (b) participation in decision making at a personally desirable level, and (c) consistency between the decision and individual preferences or values. None of the first generation IDM studies addressed all three outcomes. By far, the most attention has been focused on the impact of IDM interventions on improvements in knowledge. Some domains are in need of more representation in knowledge measures, for example, disagreement among experts whether men should be screened for prostate cancer. Two other weaknesses in knowledge measures included the use of a measure at pretest and posttest with no more than 2 weeks between them, possibly creating a testing effect; and analyses that treated individual items separately. On the other hand, careful developmental work on a knowledge scale during the planning phase of a randomized trial has yielded both a useful measure for prostate cancer screening and a model for measure development and testing (58).

Screening, intention regarding screening, or both also were measured frequently and in relatively consistent ways, often by objective indicators in the trials. Intention measures could benefit from standardization informed by reference to recommended time frames and wording (82,89). Screening, a desirable outcome for cervical, breast, and colorectal cancer, may have served as a marker for decisions having been made in the case of prostate cancer, but alone it does not signify an "informed" or "shared" decision. Authors did not typically describe their perspective on the measure, although in some instances, when screening was measured, the author described knowledge as the primary outcome. Further clarification of the role of this variable would be useful.

Although many studies assessed the respondent's preferred role in decision making, few reports of patient–provider discussion of screening were framed in parallel to role preference. Only one of these studies reported the relationship between preferred and actual role, which is necessary to fully assess the extent to which an individual participated in the decision at a level personally desired. Problems with the conceptualization and measurement of role preference discussed here need to be resolved first, however. More emphasis on measures of satisfaction with the decision-making process would be appropriate; none of the studies in this review reported using a measure for this construct, however.

Measurement of the consistency between personal values and a screening decision was even rarer in the studies that were reviewed. For circumstances in which there is considerable uncertainty about the potential benefits of an action (e.g., PSA testing), personal values are a particularly important reference point for decision making. The inclusion of values as a component of IDM reflects a shift away from paternalistic models of decision making; if there is no "right" action, the decision should reflect what is most important to the decision maker. Assuming adequate knowledge of the risks, limitations, and potential benefits of screening, the consistency between an individual's personal values and a decision may be seen as an indicator of the quality of decision making (3). It is assumed that a "high-quality" decision will increase patient satisfaction, adherence to the chosen course of action, and acceptance of the ultimate consequences of that action (as a high-quality decision may still result in clinically "poor" outcomes). Perhaps the relative lack of attention to values in the existing literature is that, in practicality, it is difficult to elucidate the specific values (other than quantity or quality of life) that are relevant to each decision, and it is also difficult to communicate them to study participants in a way that is accessible and relevant.

Utility assessment is one method, but it requires a level of numeracy that may not be reasonable to expect from all audiences. O'Connor and associates (67) proposed an alternative method to assess values relevant to decision making. In her studies, which are not among the studies reviewed here, participants engage in values clarification exercises, which involve consideration of the advantages and disadvantages of a particular course of action, and subsequent ranking or weighting of the importance of each. This approach is consistent with the foundation of much of O'Connor's work in Janis and Mann's theory of decision making, and contemporary use of decisional balance measures now associated with other models, for example, the Transtheoretical Model of Behavior Change (69,90).

In many cases, the psychometric properties of the instruments used to measure these constructs were either not reported or nonexistent. The most frequently reported characteristic of measurement instruments was internal reliability coefficients, followed by test–retest reliability as a distant second. We looked for "ceiling effects," based on reported scores and measures of variance. We did not identify this problem in the samples represented in the included studies, although it could occur in other samples and should be considered in pretesting. One study where a skewed distribution may have explained the failure to find an effect inexplicably omitted the measure from the discussion where this possibility would have become part of the accessible knowledge base (21).

Validity of measures also was reported infrequently. Content validity, established by expert review, was reported for some measures, typically constructs associated with the HBM and knowledge measures. Construct validity, criterion validity, and discriminant validity also were reported for a few well established measures. Authors did cite prior work on instrument development and testing—citing the psychometric properties of the instruments in secondary references. It was unusual to find a comment on the extent to which those psychometric properties might have been affected by use in settings, populations, or decision-making contexts that differed substantially from those in which the instrument was initially developed and tested or to

find evidence of retesting the properties of the adapted measure in our study. Data on interitem correlation, Cronbach's alpha, were sometimes reported, but authors rarely provided other evidence of validity that might have been available in their data, such as information about the correlation among measures.

The problem of conceptual overlap of measures is also evident from this review. Some of the instruments used to measure IDM constructs contain items that may be indistinguishable to the respondent. For example, an item on the Decisional Conflict Scale (69,79,86,86) asks respondents if they feel confident and "can handle unwanted pressure from others in making my choice" with responses on a 5-point Likert scale. The Decision Self-Efficacy Scale (79,81) contains a similar item asking respondents to indicate extent to which they agree or disagree with the statement, "I am making this choice without pressure from others." There are similar problems with measures of other constructs. This conceptual overlap presents numerous difficulties, including the analytic problem of collinearity and the practical issue of respondent burden. We also need to consider the relevance of decisional conflict to IDM interventions for cancer screening, and the timing of such measures. If only the Uncertainty subscale is employed, this is a little different from decision self-efficacy—and most likely relevant to PSA testing. The concept of regret also has received attention, and current thinking among social psychologists is that "anticipatory regret" plays a role in behavioral decision making only when it may influence decision making—that is, when postdecisional regret is made salient to decision makers before or when they make the decision (88).

Another issue that requires attention in the measurement of IDM constructs is the importance of using measures that are appropriate for the populations in which the intervention occurs. There has been a call for the development of IDM interventions for diverse populations, such as those who are non-White, older, and medically underserved (3). Yet much of the existing work on IDM interventions has been conducted among predominantly White samples with relatively high levels of education. Given the importance of targeting interventions to underserved audiences, more work is needed to develop measures that are -appropriate for low literacy populations, as, for example, O'Conner and associates have done for decisional self-efficacy (79). More cross-cultural research is needed to better understand the influence of culture on outcomes for IDM, so that appropriate measures can be developed. For example, some cultures value independence in decision making, whereas others consider respect for authorities (i.e., the physician) paramount. Obviously, this could affect the ways in which population subgroups respond to IDM measures (e.g., role preference). Similarly, values placed on potential outcomes are likely to vary by culture and thus require further exploration and documentation.

Last, we recognize that new behavioral foci for cancer screening IDM are emerging. With increasing attention to the potential of widespread testing for HPV, studies of IDM for HPV testing will be increasingly important (91,92). Given the key role of this type of study in public health policy, the research can be most useful if it builds on past measurement studies and follows the guidelines provided here.

RECOMMENDATIONS

Development and use of validated measures to assess the primary outcomes of IDM interventions is critical to advancing the field of decision making in cancer screening. One step toward this goal is greater use and application of theoretical frameworks to the design of IDM studies. A minority of studies included in this review cited a conceptual framework or theoretical model(s) on which the intervention and evaluation were based, as documented elsewhere in this issue (93).

Use of theoretical frameworks is important both for developing interventions and designing the evaluation. In terms of intervention development, theory can suggest intervention strategies and articulate the assumptions behind the choice of intervention strategies and components. In planning evaluations, theory can (a) pinpoint primary and secondary outcomes, (b) identify potential mediating or moderating factors to assess, (c) provide conceptual clarity, improve operational precision, (d) determine timing of data collection, and (e) dictate analytic strategies. This would help overcome the heterogeneous use of variables in analysis observed in this review. The analytic framework presented in the Task Force review (3) could serve as a foundation. Moreover, we believe that studies of IDM could benefit from a greater exploration of the theoretical and empirical work in other relevant fields, such as economics, decision analysis, and clinical psychology as discussed in another article in this issue (93).

More use of primary sources for measures, now increasingly available on Web sites, in conjunction with a guide for selecting measurement instruments such as one created by DeVellis (50) would help investigators make clearer, more informed choices about measures. For example, the DeVellis guide, includes considerations such as feasibility of use with a particular population, selection of multiple indicators of important concepts, and comments about pitfalls in interpreting psychometric characteristics presented by other authors. Web sites are particularly important sources, because users' manuals and other information that may not have been included in journal articles; journal articles and unpublished guides for using measures; adaptations; and, most important, the instruments themselves can all be posted. Several examples of Web sites have been included in previous sections.

Studies in progress and completed studies can, as suggested in our comments about missed opportunities to derive useful information about the performance and characteristics of measures, contribute to the information base on measures. Analyses such as those used for the Prostate Cancer Screening Education Study trial knowledge measure development paper (58) provide useful ideas, as do resources such as the DeVellis guide (50).

Adherence to reporting guidelines and requirements such as Consolidated Standards of Reporting Trials (CONSORT) (94) and Transparent Reporting of Evaluations with Non-randomized Designs (TREND) (95) would improve the clarity and information provided on measures. Looking specifically at the standards that pertain to measures in these guides, we see several that would have improved our ability to identify characteristics of measures—had they been followed more often. These in-

clude (a) which of the multiple outcome measures is primary, and which are secondary, (b) what methods were used to enhance the quality of measurements (e.g., multiple observations, training of assessors), and (c) were those assessing the outcomes blinded as to respondents' group assignment? Implied in these standards also is clear designation of the timing and method of data collection for each measure. We urge the consistent use of these guidelines as policy in all journals, because CONSORT, in particular, has had an enormous impact on the completeness, clarity, and consistent reporting of intervention trials in medical journals, and their application in articles in this article made coding far easier. Further reporting categories suggested here are (a) the construct the measure was intended to represent; (b) the role of the construct in the analytic framework of the study; (c) a summary of information already known about the measure and its likely generalizabiilty to this study—in particular, when the wording of the measure has been changed, items removed, and other adaptations have been made; (d) where the full measure can be found, if it is not reproduced in the publication; (e) information about the measure's performance in our study; and (f) discussion of how measurement performance may have influenced study results.

Finally, this work can only be accomplished if its importance is recognized by funding agencies, researchers, and peer-reviewed journal editors and reviewers. Without the tools for conducting rigorous evaluations, resources spent on large evaluation studies are not yet a sound investment. However, funding for methods development lags behind that of other research. Continual improvement of data collection measures is an important goal for all types of research but is particularly important at this stage of development in the field of IDM, where the number of studies is likely to increase rapidly, given a growing desire on the part of patients for involvement in decision making, increased access to screening tests with uncertain medical benefit (e.g., PSA), and a shift away from paternalistic models of care (3,7). Increased opportunities and infrastructure for collaborations, such as the Cancer Prevention and Control Network, helped to make this article possible and could be a model for future work in this area.

REFERENCES

(1) Braddock CH, Edwards KA, Hasenberg NM, Laidley TL, Levinson W: Informed decision making in outpatient practice: Time to get back to basics. *Journal of the American Medical Association.* 1999, *282:*2313–2320.

(2) Emanuel EJ, Emanuel LL: Four models of the physician–patient relationship [review]. *Journal of the American Medical Association.* 1992, *267:*2221–2226.

(3) Briss P, Rimer B, Reilley B, et al.: Promoting informed decisions about cancer screening in communities and healthcare systems. *American Journal of Preventive Medicine.* 2004, *26:*67–80.

(4) O'Connor AM, Fiset V, DeGrasse C, et al.: Decision aids for patients considering options affecting cancer outcomes: Evidence of efficacy and policy implications. *Journal of the National Cancer Institute Monographs.* 1999, 67–80.

(5) O'Connor AM, Stacey D, Entwistle V, et al.: Decision aids for people facing health treatment or screening decisions. *Cochrane Database of Systematic Reviews.* 2003, 86, CD001431.

(6) Rimer BK, Briss PA, Zeller PK, Chan EC, Woolf SH: Informed decision making: What is its role in cancer screening? *Cancer.* 2004, *101:*1214–1228.

(7) Sheridan SL, Harris RP, Woolf SH: U.S. Preventive Services Task Force: Shared decision making about screening and chemoprevention. A suggested approach from the U.S. Preventive Services Task Force. *American Journal of Preventive Medicine.* 2004, *26:*56–66.

(8) Elwyn G, Edwards A, Eccles M, Rovner D: Decision analysis in patient care [review]. *Lancet*. 2001, *358:*571–574.

(9) Entwistle VA, Sheldon TA, Sowden A, Watt IS: Evidence-informed patient choice. Practical issues of involving patients in decisions about health care technologies. *International Journal of Technology Assessment in Health Care.* 1998, *14:*212–225.

(10) Rimer BK, Halabi S, Sugg SC, et al.: Effects of a mammography decision-making intervention at 12 and 24 months. *American Journal of Preventive Medicine.* 2002, *22:*247–257.

(11) Wolf AM, Becker DM: Cancer screening and informed patient discussions. Truth and consequences. *Archives of Internal Medicine.* 1996, *156:*1069–1072.

(12) Flood AB, Wennberg JE, Nease RF, et al.: The importance of patient preference in the decision to screen for prostate cancer [Study 1 of 2: Free-PSA Screening Study]. *Journal of General Internal Medicine.* 1996, *11:*342–349.

(13) Flood AB, Wennberg JE, Nease RF, et al.: The importance of patient preference in the decision to screen for prostate cancer [Study 2 of 2: Scheduled Clinic Study]. *Journal of General Internal Medicine.* 1996, *11:*342–349.

(14) Bridge PD, Berry-Bobovski L, Bridge TJ, Gallagher RE: Evaluation of a preparatory community-based prostate health education program. *Journal of Cancer Education.* 2002, *17:*101–105.

(15) Bridge PD, Berry-Bobovski LC, Gallagher RE: Promoting informed decision making: Evaluating a community-based prostate health awareness program. *Journal of Cancer Education.* 2004, *19:*174–179.

(16) Frosch DL, Kaplan RM, Felitti V: The evaluation of two methods to facilitate shared decision making for men considering the prostate-specific antigen test. *Journal of General Internal Medicine.* 2001, *16:*391–398.

(17) Frosch DL, Kaplan RM, Felitti VJ: A randomized controlled trial comparing Internet and video to facilitate patient education for men considering the prostate specific antigen test. *Journal of General Internal Medicine.* 2003, *18:*781–787.

(18) Gattellari M, Ward JE: Does evidence-based information about screening for prostate cancer enhance consumer decision-making? A randomised controlled trial. *Journal of Medical Screening.* 2003, *10:*27–39.

(19) Gattellari M, Ward JE: A community-based randomised controlled trial of three different educational resources for men about prostate cancer screening. *Patient Education and Counseling*. 2005, *57:*168–182.

(20) Partin MR, Nelson D, Radosevich D, et al.: Randomized trial examining the effect of two prostate cancer screening educational interventions on patient knowledge, preferences, and behaviors. *Journal of General Internal Medicine.* 2004, *19:*835–842.

(21) Volk RJ, Spann SJ, Cass AR, Hawley ST: Patient education for informed decision making about prostate cancer screening: A randomized controlled trial with 1-year follow-up. *Annals of Family Medicine.* 2003, *1:*22–28.

(22) Wilt TJ, Paul J, Murdoch M, et al.: Educating men about prostate cancer screening. A randomized trial of a mailed pamphlet. *Effective Clinical Practice.* 2001, *4:*112–120.

(23) Sheridan SL, Felix K, Pignone MP, Lewis CL: Information needs of men regarding prostate cancer screening and the effect of a brief decision aid. *Patient Education and Counseling.* 2004, *54:*345–351.

(24) Schapira MM, VanRuiswyk J: The effect of an illustrated pamphlet decision-aid on the use of prostate cancer screening tests. *Journal of Family Practice.* 2000, *49:*418–424.

(25) Weinrich SP, Seger R, Miller BL, et al.: Knowledge of the limitations associated with prostate cancer screening among low-income men. *Cancer Nursing.* 2004, *27:*442–453.

(26) Adams JM: The effect of a health information intervention on knowledge and health practices regarding colorectal cancer in African-Americans, UMI Order PUZ9630142. Unpublished Doctoral Dissertation, Texas Woman's University: 1996.

(27) Braun KL, Fong M, Kaanoi ME, Kamaka ML, Gotay CC: Testing a culturally appropriate, theory-based intervention to improve colorectal cancer screening among Native Hawaiians. *Preventive Medicine.* 2005, *40:*619–627.

(28) Campbell MK, James A, Hudson MA, et al.: Improving multiple behaviors for colorectal cancer prevention among African American church members. *Health Psychology.* 2004, *23:*492–502.

(29) Dolan JG, Frisina S: Randomized controlled trial of a patient decision aid for colorectal cancer screening. *Medical Decision Making.* 2002, *22:*125–139.

(30) Wolf AM, Schorling JB: Does informed consent alter elderly patients' preferences for colorectal cancer screening? Results of a randomized trial. *Journal of General Internal Medicine.* 2000, *15:*24–30.

(31) Lewis CL, Pignone MP, Sheridan SL, Downs SM, Kinsinger LS: A randomized trial of three videos that differ in the framing of information about mammography in women 40 to 49 years old. *Journal of General Internal Medicine.* 2003, *18:*875–883.

(32) Rimer BK, Halabi S, Sugg SC, et al.: The short-term impact of tailored mammography decision-making interventions. *Patient Education and Counseling.* 2001, *43:*269–285.

(33) Clark MA, Rakowski W, Ehrich B, et al.: Stages of adopting regular screening mammography: Do women differ in decisional balance within stages? *Journal of Health Psychology.* 1998, *3:*491–506.

(34) Davison BJ, Kirk P, Degner LF, Hassard TH: Information and patient participation in screening for prostate cancer. *Patient Education and Counseling.* 1999, *37:*255–263.

(35) Wolf AM, Schorling JB: Preferences of elderly men for prostate-specific antigen screening and the impact of informed consent. *Journals of Gerontology Series A:Biological and Medical Sciences.* 1998, *53:*M195–M200.

(36) Vadaparampil ST, Jacobsen PB, Kash K, et al.: Factors predicting prostate specific antigen testing among first-degree relatives of prostate cancer patients. *Cancer Epidemiology, Biomarkers and Prevention.* 2004, *13:*753–758.

(37) Messina CR, Lane DS, Grimson R: Colorectal cancer screening attitudes and practices preferences for decision making. *American Journal of Preventive Medicine.* 2005, *28:*439–446.

(38) Watson EJG: Decision-making style, motivation, and use of papnicolaou screening tests in women age 35 and older, ProQuest No. ATT 9322004. Unpublished Doctoral Dissertation, University of South Carolina: 1993.

(39) Wackerbarth SB, Peters JC, Haist SA: "Do we really need all that equipment?": Factors influencing colorectal cancer screening decisions. *Qualitative Health Research.* 2005, *15:*539–554.

(40) Davey C, White V, Gattellari M, Ward JE: Reconciling population benefits and women's individual autonomy in mammographic screening: In-depth interviews to explore women's views about 'informed choice.' *Australian and New Zealand Journal of Public Health.* 2005, *29:*69–77.

(41) Ruthman JL, Ferrans CE: Efficacy of a video for teaching patients about prostate cancer screening and treatment. *American Journal of Health Promotion.* 2004, *18:*292–295.

(42) Nekhlyudov L, Li R, Fletcher SW: Information and involvement preferences of women in their 40s before their first screening mammogram. *Archives of Internal Medicine.* 2005, *165:*1370–1374.

(43) Volk RJ, Cantor SB, Spann SJ, et al.: Preferences of husbands and wives for prostate cancer screening. *Archives of Family Medicine.* 1997, *6:*72–76.

(44) Volk RJ, Cantor SB, Cass AR, et al.: Preferences of husbands and wives for outcomes of prostate cancer screening and treatment. *Journal of General Internal Medicine.* 2004, *19:*339–348.

(45) Pignone M, Harris R, Kinsinger L: Videotape-based decision aid for colon cancer screening. A randomized, controlled trial. *Annals of Internal Medicine.* 2000, *133:*761–769.

(46) Pignone M, Bucholtz D, Harris R: Patient preferences for colon cancer screening. *Journal of General Internal Medicine.* 1999, *14:*432–437.

(47) Fowler DA: Claiming health: mammography screening decision-making of older, urban African-American women, ProQuest No. AAT 3082875. Unpublished Doctoral Dissertation, Rush University College of Nursing: 2003.

(48) Nunnally JC: *Psychometric Theory* (2nd Ed.). New York: McGraw-Hill, 1978.

(49) DeVellis RF: *Scale Development: Theory and Applications*. Newbury Park, CA: Sage, 1991.

(50) DeVellis RF: A consumer's guide to finding, evaluating, and reporting on measurement instruments. *Arthritis Care and Research.* 1996, *9:*239–245.

(51) Finney Rutten LJ, Meissner HI, Breen N, Vernon SW, Rimer BK: Factors associated with men's use of prostate-specific antigen screening: Evidence from Health Information National Trends Survey. *Preventive Medicine.* 2005, *40:*461–468.

(52) Gattellari M, Ward JE: A community study using specified and unspecified scenarios to investigate men's views about PSA screening. *Health Expectations.* 2004, *7:*274–289.

(53) O'Dell KJ, Volk RJ, Cass AR, Spann SJ: Screening for prostate cancer with the prostate-specific antigen test: Are patients making informed decisions? *Journal of Family Practice.* 1999, *48:*682–688.

(54) Volk RJ, Cass AR, Spann SJ: A randomized controlled trial of shared decision making for prostate cancer screening. *Archives of Family Medicine.* 1999, *8:*333–340.

(55) Wolf AMD, Nasser JF, Wolf AM, Schorling JB: The impact of informed consent on patient interest in prostate-specific antigen screening. *Archives of Internal Medicine.* 1996, *156:*1333–1336.
(56) Katz ML, James AS, Pignone MP, et al.: Colorectal cancer screening among African American church members: A qualitative and quantitative study of patient–provider communication. *BMC Public Health.* 2004, *4:*62.
(57) Chan EC, Sulmasy DP: What should men know about prostate-specific antigen screening before giving informed consent? *American Journal of Medicine.* 1998, *105:*266–274.
(58) Radosevich DM, Partin MR, Nugent S, et al.: Measuring patient knowledge of the risks and benefits of prostate cancer screening. *Patient Education and Counseling.* 2004, *54:*143–152.
(59) Rosenstock IM: Historical origins of the Health Belief Model. *Health Education Monographs.* 1974, *2:*328–335.
(60) Rosenstock IM, Strecher VJ, Becker MH: Social Learning Theory and the Health Belief Model. *Health Education Quarterly.* 1988, *15:*175–183.
(61) Spielberger CD, Gorsuch RL: *State-Trait Anxiety Inventory for Adults: Sampler Set: Manual, Test, Scoring Key.* Redwood City, CA: Mind Garden, 1983.
(62) Consedine NS, Magai C, Krivoshekova YS, Ryzewicz L, Neugut AI: Fear, anxiety, worry, and breast cancer screening behavior: A critical review. *Cancer Epidemiology, Biomarkers and Prevention.* 2004, *13:*501–510.
(63) Hay JL, Buckley TR, Ostroff JS: The role of cancer worry in cancer screening: A theoretical and empirical review of the literature. *Psycho-Oncology.* 2005, *14:*517–534.
(64) Champion VL: Instrument refinement for breast cancer screening behaviors. *Nursing Research.* 1993, *42:*139–143.
(65) Stein JA, Fox SA, Murata PJ, Morisky DE: Mammography usage and the Health Belief Model. *Health Education Quarterly.* 1992, *19:*447–462.
(66) Lerman C, Lustbader E, Rimer B, et al.: Effects of individualized breast cancer risk counseling: A randomized trial. *Journal of the National Cancer Institute.* 1995, *87:*286–292.
(67) O'Connor AM, Tugwell P, Wells GA, et al.: A decision aid for women considering hormone therapy after menopause: Decision support framework and evaluation. *Patient Education and Counseling.* 1998, *33:*267–279.
(68) Janis IL, Mann L: *Decision Making: A Psychological Analysis of Conflict, Choice, and Commitment*. New York: Free Press, 1977.
(69) Rakowski W, Dube CE, Marcus BH, et al.: Assessing elements of women's decisions about mammography. *Health Psychology.* 1992, *11:*111–118.
(70) Rakowski W, Andersen MR, Stoddard AM, et al.: Confirmatory analysis of opinions regarding the pros and cons of mammography. *Health Psychology.* 1997, *16:*433–441.
(71) Rakowski W, Clark MA, Pearlman DN, et al.: Integrating pros and cons for mammography and Pap testing: Extending the construct of decisional balance to two behaviors. *Preventive Medicine.* 1997, *26:*664–673.
(72) Degner LF, Sloan JA: Decision making during serious illness: What role do patients really want to play? *Journal of Clinical Epidemiology.* 1992, *45:*941–950.
(73) Degner LF, Kristjanson LJ, Bowman D, et al.: Information needs and decisional preferences in women with breast cancer. *JAMA: Journal of the American Medical Association.* 1997, *277:*1485–1492.
(74) Degner LF, Sloan JA, Venkatesh P: The Control Preferences Scale. *Canadian Journal of Nursing Research*. 1997, *29:*21–43.
(75) Festinger L: *Conflict, Decision, and Dissonance*. Palo Alto, CA: Stanford University Press, 1964.
(76) Davey HM, Armstrong BK, Butow PN: An exploratory study of cancer patients' views on doctor-provided and independent written prognostic information. *Patient Education and Counseling.* 2005, *56:*349–355.
(77) Deber RB, Kraetschmer N, Irvine J: What role do patients wish to play in treatment decision making? *Archives of Internal Medicine.* 1996, *156:*1414–1420.
(78) Ende J, Kazis L, Ash A, Moskowitz MA: Measuring patients' desire for autonomy: Decision making and information-seeking preferences among medical patients. *Journal of General Internal Medicine.* 1989, *4:*23–30.
(79) Ottawa Health Research Institute: *Patient Decision Aids: Evaluation Measures.* Retrieved April 7, 2006 from http://decisionaid.ohri.ca/eval.html
(80) O'Connor AM, Perrault DJ: Importance of physician's role highlighted in survey of women's breast screening practices. *Canadian Journal of Public Health Revue Canadienne de Sante Publique.* 1995, *86:*42–45.
(81) Bunn H, O'Connor A: Validation of client decision-making instruments in the context of psychiatry. *Canadian Journal of Nursing Research.* 1996, *28:*13–27.
(82) Ajzen I: *Constructing a TpB [Theory of Planned Behavior] Questionnaire: Conceptual and Methodological Considerations.* Retrieved April 7, 2006 from http://www.people.umass.edu/aizen/pdf/tpb.measurement.pdf
(83) Holmes-Rovner M, Kroll J, Schmitt N, et al.: Patient satisfaction with health care decisions: The Satisfaction with Decision Scale. *Medical Decision Making.* 1996, *16:*58–64.
(84) Wills CE, Holmes-Rovner M: Patient comprehension of information for shared treatment decision making: State of the art and future directions. *Patient Education and Counseling.* 2003, *50:*285–290.
(85) Barry MJ, Cherkin DC, YuChiao C, Fowler FJ, Skates S: A randomized trial of a multimedia shared decision-making program for men facing a treatment decision for benign prostatic hyperplasia. *Disease Management and Clinical Outcomes.* 1997, *1:*5–14.
(86) O'Connor AM: Validation of a decisional conflict scale. *Medical Decision Making.* 1995, *15:*25–30.
(87) Brehm JW, Cohen AR: *Explorations in Cognitive Dissonance*. New York: Wiley, 1962.
(88) Zeelenberg M: Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making.* 1999, *12:*93–106.
(89) Ajzen I, Brown TC, Carvajal F: Explaining the discrepancy between intentions and actions: The case of hypothetical bias in contingent valuation. *Personality and Social Psychology Bulletin.* 2004, *30:*1108–1121.
(90) Prochaska JO, Velicer WF, Rossi JS, et al.: Stages of change and decisional balance for 12 problem behaviors. *Health Psychology.* 1994, *13:*39–46.
(91) Anhang R, Goodman A, Goldie SJ: HPV communication: Review of existing research and recommendations for patient education. *CA: A Cancer Journal for Clinicians.* 2004, *54:*248–259.
(92) Kim JJ, Wright TC, Goldie SJ: Cost-effectiveness of human papillomavirus DNA testing in the United Kingdom, The Neth-

erlands, France, and Italy. *Journal of the National Cancer Institute.* 2005, *97:*888–895.

(93) Bowen DJ, Allen JD, Vu T, et al.: Theoretical foundations for interventions designed to promote IDM for cancer screening. *Annals of Behavioral Medicine.* 2006. *32:*202–210. [this issue]

(94) Altman DG, Schulz KF, Moher D, et al.: The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Annals of Internal Medicine.* 2001, *134:*663–694.

(95) Des Jarlais DC, Lyles C, Crepaz N: Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health.* 2004, *94:*361–366.