Clinicians' Perceptions of the Benefits and Harms of Prostate and Colorectal Cancer Screening

Emily A. Elstad, PhD, Anne Sutkowi-Hemstreet, MPH, Stacey L. Sheridan, MD, MPH, Maihan Vu, DrPH, Russell Harris, MD, MPH, Valerie F. Reyna, PhD, Christine Rini, PhD, Jo Anne Earp, ScD, Noel T. Brewer, PhD

Background. *Clinicians' perceptions of screening benefits and harms influence their recommendations, which in turn shape patients' screening decisions. We sought to understand clinicians' perceptions of the benefits and harms of cancer screening by comparing 2 screening tests that differ in their balance of potential benefits to harms: colonoscopy, which results in net benefit for many adults, and prostate-specific antigen (PSA) testing, which may do more harm than good.* ***Methods.*** *In this cross-sectional study, 126 clinicians at 24 family/internal medicine practices completed surveys in which they listed and rated the magnitude of colonoscopy and PSA testing benefits and harms for a hypothetical 70-year-old male patient and then estimated the likelihood that these tests would cause harm and lengthen the life of 100 similar men in the next 10 years. We tested the hypothesis that the availability heuristic would explain the association of screening test to perceived likelihood of benefit/harm and a competing hypothesis that clinicians' gist of screening tests as good or bad would mediate this association.* ***Results.*** *Clinicians perceived PSA testing to have a greater likelihood of harm and a lower likelihood of lengthening life relative to colonoscopy. Consistent with our gist hypothesis, these associations were mediated by clinicians' gist of screening (balance of perceived benefits to perceived harms).* ***Limitations.*** *Generalizability beyond academic clinicians remains to be established.* ***Conclusions.*** *Targeting clinicians' gist of screening, for example through graphical displays that allow clinicians to make gist-based relative magnitude comparisons, may influence their risk perception and possibly reduce overrecommendation of screening.* ***Key words:*** *prostate cancer screening; colorectal cancer screening; PSA test; colonoscopy; risk perception; clinicians.* ***(Med Decis Making 2015;35:467–476)***

Cancer screening poses both potential benefits and potential harms to patients. It can lengthen life and increase quality of life by reducing cancer-related morbidity. However, overuse of cancer screening can result in harms from the screening procedure itself and from overdiagnosis and unnecessary follow-up and treatment.[1] Clinicians' recommendations are instrumental in shaping patients' screening decisions,[2,3] yet we know little about clinicians' perceptions of screening benefits and harms or how they arrive at their perceptions of the likelihood of benefit or harm from screening. These kinds of perceptions and likelihood judgments are a useful focus of research because they play an important role in theories of decision making[4] and health behavior.[5]

To understand how clinicians formulate these perceptions and likelihood judgments, it is helpful first to know that people often do not use calculated, rational decision strategies[6] but instead rely on quick, intuitive, automatic strategies, sometimes called *heuristics*, to make decisions under uncertainty.[7] Laypeople and experts, including clinicians, frequently rely up heuristics,[8–10] more so as expertise increases.[11–14] One such heuristic is the availability heuristic, whereby people estimate the likelihood of future events based on the ease with which they can call to mind instances of such events.[15] For example, patients[16] and physicians[17,18] tend to overestimate the likelihood of a disease if they can more easily recall details about it. Researchers frequently operationalize availability of information in 2 ways: as the number of instances of a particular type of

DOI: 10.1177/0272989X15569780

information that participants can recall[15,19,20] and as the subjective ease of recalling those instances.[19,21,22] By this reasoning, clinicians who recall more benefits of screening with greater ease may also perceive a greater likelihood of benefit from screening.

Another possibility is that clinicians may perceive the benefits and harms of screening as a gestalt, again more so as expertise increases.[12,13,23,24] According to fuzzy trace theory, memories of precise, verbatim information (e.g., specific probabilities such as ''my patient has a 2 in 1000 chance of experiencing a harm'') fade quickly over time; more enduring is gist memory, or the bottom-line meaning ascribed to an event (e.g., ''my patient's chance of experiencing harm is remote'').[25] Individuals generally rely on gist information, even when they can remember verbatim information,[25] and they may base likelihood estimates on gist impressions rather than disease prevalence.[26] Clinicians' gist of screening may manifest as an overall impression of net benefit or harm, which takes into account both the number of benefits and harms and the magnitude of those benefits and harms.[27] Thus, if clinicians have a negative gist of screening (i.e., they ascribe greater total magnitude to harms than benefits), they may judge the likelihood of harm from screening to be higher. Similarly, if they have a positive gist of screening, they may judge the likelihood of benefits from screening to be higher.

Received 6 April 2014 from the Department of Health Behavior, University of North Carolina, Chapel Hill, NC (EAE, CR, JAE, NTB); Cecil B. Sheps Center for Health Services Research, University of North Carolina, Chapel Hill, NC (EAE, ASH, SLS, MV, RH); Center for Health Promotion and Disease Prevention, University of North Carolina, Chapel Hill, NC (SLS, MV); Health Care and Prevention Program, University of North Carolina, Chapel Hill, NC (SLS, RH); Division of General Medicine and Clinical Epidemiology, University of North Carolina, Chapel Hill, NC (SLS, RH); College of Human Ecology, Cornell University, Ithaca, NY (VFR); and Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC (CR, NTB). This research was partially supported by a National Service Research Award Pre-Doctoral Traineeship from the Agency for Health Care Research and Quality (AHRQ) sponsored by the Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, grant 5T32 HS000032; and by AHRQ grant P01-HS21133-02. Financial support for this study was provided in part by a grant from the Agency for Healthcare Research and Quality. The funding agreement ensured the authors' independence in designing the study, interpreting the data, and writing and publishing the report. Revision accepted for publication 19 December 2014.

Address correspondence to Noel T. Brewer, Department of Health Behavior, Gillings School of Global Public Health, University of North Carolina, 325 Rosenau Hall, CB #7440, Chapel Hill, NC 27599-7440; e-mail: ntb@unc.edu.

To better understand clinicians' perceptions of screening benefits and harms, the present study compared clinicians' perceptions of 2 screening tests that vary in their balance of benefits and harms. Colonoscopy is a high-efficacy screening test that has been shown to result in net benefit in adults ages 50–75 and reduces colorectal cancer mortality,[28,29] and national recommendations suggest its use for that age group.[28–30] The second screening test we chose was prostate-specific antigen (PSA) testing, which has been shown to lead to net harm, and national screening recommendations discourage the test[31] or recommend it only conditionally.[32,33]

The study had 2 distinct but complementary aims. First, we sought to describe clinicians' perceptions of the specific benefits and harms of our 2 chosen screening services, including the number and perceived magnitude of benefits and harms they could call to mind. Second, we sought to understand how clinicians arrive at their perceptions of the likelihood of life lengthened or harm from screening. We predicted that clinicians would perceive the likelihood of harm to be greater and likelihood of life lengthened to be lower for a screening test with harms that outweigh benefits (PSA testing) relative to a screening test with benefits that outweigh harms for many adults (colonoscopy). We had 2 competing hypotheses about mediators of this association. Our *availability hypothesis* was that availability would explain the association of screening test to perceived likelihood of benefit or harm, consistent with the availability heuristic. We predicted that clinicians would list more harms and fewer benefits for PSA testing relative to colonoscopy and that they would have less difficulty recalling PSA testing harms relative to colonoscopy harms. We further predicted that these variables would mediate the association of screening test (PSA v. colonoscopy) to clinicians' perceived likelihood of harm and life lengthened. Our *fuzzy trace hypothesis* was that clinicians' gist of screening tests as good or bad would mediate the association between screening test and clinicians' likelihood perceptions. We predicted that clinicians would perceive the likelihood of harm to be greater and likelihood of life lengthened to be lower for PSA relative to colonoscopy if their gist of colonoscopy was more positive than their gist of PSA testing.

METHODS

Participants

Eligible participants were clinicians at 24 family medicine or internal medicine practices in a North

Carolina university-affiliated, practice-based research network. In fall 2012, practices in the network employed 155 practicing clinicians: 127 medical doctors, 3 doctors of osteopathic medicine, 16 physician assistants, and 12 nurse practitioners. We excluded clinicians who did not have their own panel of patients.

Procedures

As part of a study of the harms of clinical preventive services, we administered a paper survey to clinicians. Practice representatives distributed the surveys to clinicians in their practices. The study packet included a $20 bill as incentive to complete and return the survey.[34] The institutional review board of the University of North Carolina approved the study protocol and materials.

The survey included 2 vignettes that held a hypothetical patient's sex, age, race, health status, and family and screening history constant but varied the screening test (PSA testing or colonoscopy). We counterbalanced the order of questions on screening tests by randomly assigning clinicians to 1 of 2 questionnaire conditions in which prostate or colorectal cancer screening vignettes and questions appeared first. The hypothetical patient for PSA testing was Mr. Morton, a 70-year-old white male with good cognitive status, no fatal disease, no family history of prostate cancer, no previous prostate findings or abnormal PSA tests, and a normal PSA test result 2 years ago. The hypothetical patient for colonoscopy was Mr. Lewis, a 70-year-old white male with good cognitive status, no fatal disease, no family history of colon cancer, no risk factors or history of polyps, and a normal colonoscopy result 10 years ago.

Measures

Outcomes

Perceived likelihood of harm. The survey measured perceived likelihood of harm with the question, ''Think of 100 healthy patients, like Mr. Morton [Mr. Lewis] age 70, whom you screen with the PSA test [colonoscopy] and find an elevated PSA of 8.0 [a 1.0-cm adenomatous polyp that is removed]. You continue to follow them for the next 10 years. Having the PSA test [colonoscopy] will lead to at least moderate physical harm at some point over the next 10 years for how many of these men?'' Response options were 0, 1–10, 11–20, 21–30, 31–40, 41–50, 51–60, 61–70, 71–80, 81–90, and 91–100 men out of 100 men.

Perceived likelihood of life lengthened. The survey measured perceived likelihood of life lengthened with the question, ''Think of 100 healthy patients, like Mr. Morton [Mr. Lewis] age 70, whom you screen with the PSA test [colonoscopy] and find an elevated PSA of 8.0 [a 1.0-cm adenomatous polyp that is removed]. You continue to follow them for the next 10 years. At the end of 10 years, how many of these men do you think will have had their lives lengthened by having had the PSA test [colonoscopy]?'' Response options were the same as for perceived likelihood of harm.

Mediators

Number of benefits and harms. We measured availability as the number of benefits [harms] of PSA testing [colonoscopy] with the question, ''List as many benefits [harms] from PSA testing [colonoscopy] as you can think of for Mr. Morton [Mr. Lewis], a 70-year-old patient.''[15,19,20] Clinicians could list up to 7 benefits [harms], a number deemed adequate by clinicians on the study team. We instructed clinicians to use only the lines they needed.

Subjective ease of recall. The survey measured subjective difficulty of recall by asking, ''On average, how difficult was it for you to come up with these harms for prostate [colorectal] cancer screening?'' The 5-point response scale ranged from *not at all* (coded 0) to *extremely* (coded 4).

Magnitude of benefit [harm]. For each benefit [harm] that clinicians listed, the survey asked them to ''indicate how large you believe that benefit [harm] would be.'' Response options were almost no benefit [harm] to patient (coded as 1), small benefit [harm], moderate benefit [harm], and large benefit [harm] (coded as 4). We calculated the magnitude of benefit [harm] as the sum of the ratings of each benefit [harm] a clinician listed.

Gist. We operationalized gist as the difference between perceived benefits of screening and perceived harms of screening, drawing upon methods used previously.[35] For each test, we summed the magnitude ratings of listed harms and separately summed the magnitude ratings of listed benefits. We then calculated gist for each test as the summed magnitude of benefits minus the summed magnitude of harms.[27] A positive gist score indicated that a clinician listed more benefits with greater magnitude than harms, whereas a negative gist score indicated that a clinician listed more harms with greater magnitude than benefits.

Data Analyses

Two researchers (E.E. and M.V.) tabulated the benefits and harms that clinicians listed for each test and then established a classification of benefits and harms. Two coders (E.E. and A.S.H.) independently categorized the harms that clinicians listed into 5 categories: physical effects, psychological effects, financial strain, opportunity costs, and hassle (i.e., sometimes unnecessary difficulties associated with complex requirements of testing and treatment). These categories were informed by the taxonomy of screening harms proposed by Harris and colleagues.[1] Interrater reliabilities for each category were good (Cohen's kappa > 0.80).

Paired *t* tests compared the mean number of PSA testing harms that clinicians listed to the mean number of PSA testing benefits. We repeated this test for colonoscopy and for the magnitude sum scores, perceived likelihood of harm, and perceived likelihood of life lengthened. We used paired *t* tests to compare mean PSA testing benefits to mean colonoscopy benefits. We repeated this analysis for harms, subjective difficulty of recall, magnitude sum scores for benefits and harms, gist, likelihood of harm, and likelihood of life lengthened. McNemar tests compared the frequency of mentions of each benefit and harm category between PSA test and colonoscopy.

Missingness on variables was 2%–5% for demographics, 10%–19% for mediators, and 6% for outcomes. Some clinicians did not list benefits or harms on the survey, possibly due to the extra burden of doing so. As a result, up to 19% of values were missing for the number and magnitude of benefits and harms and gist measures. Thus, gist had the most missing data (19%), as we calculated this variable from other variables. We used multiple imputation with the expectation-maximization algorithm to impute missing data and reduce bias. This algorithm computes missing observations given the observed data and replaces missing observations with the conditional mean based on the regression equations.[36] Based on exploratory analyses, we determined our data to be missing at random as required by multiple imputation.[37] Auxiliary variables in the imputation included all the variables in the mediation analyses. We imputed 300 datasets and ran all models for each dataset.[38] SAS was then used to average parameter estimates across each analysis and calculate standard errors that combined variability within and between data sets.

The main outcome measures were clinicians' perceived likelihood of harm and life lengthened from screening. We used generalized estimating equations that accounted for repeated measurements to examine whether perceived likelihood of harm differed by screening test (PSA v. colonoscopy). We repeated this analysis to assess the association between screening test and perceived likelihood of lengthening life. In separate models, we then tested several potential mediators of these associations: number of harms and benefits and subjective difficulty of recall (availability hypothesis); gist (fuzzy trace hypothesis); and additional gist components (perceived magnitude of benefit and perceived magnitude of harm). We used a causal steps approach to mediation.[39] Consistent with that approach, we tested the following: the associations described above; whether screening test predicted potential mediator variables; whether mediator variables predicted likelihood estimates statistically controlling for screening test; and whether the effect of screening test on likelihood estimates was attenuated after controlling for the effect of gist and gist components on likelihood estimates in separate models.[29] In each model, we controlled for the order in which clinicians viewed questions on each screening test. We conducted Sobel tests to establish whether reductions in the association in step 4 were attributable to the mediators. We conducted all analyses in SAS 9.2[40] using 2-tailed tests and a critical alpha of .05. The funding source for this study (the Agency for Health Care Research and Quality) played no role in this study.

RESULTS

A total of 126 clinicians returned the survey (80% response rate). Respondents were primarily male (62%) and physicians (79%). Seventy-six percent of participants were white, 11% Asian, and 10% black or African American. Participants were 45 years old on average, and mean years in medical practice was 15.

PSA Testing

The benefits of PSA testing that clinicians most frequently mentioned were early detection and treatment (72%) and psychological effects (e.g., peace of mind) (37%) (Table 1). The most frequently listed harms were unnecessary treatment (56%), psychological effects (e.g., anxiety) (53%), and follow-up (47%). Many clinicians listed at least 1 physical harm of PSA testing (70%) and many listed at least 1 psychological harm (68%) or hassle (56%). However, fewer clinicians recognized financial strain (13%). No clinicians listed opportunity costs of

Table 1 Frequency of Mentions and Mean Magnitude of PSA Testing and Colonoscopy Benefits and Harms

	Listed for PSA, %	Listed for Colonoscopy, %	Magnitude Rating for PSA, $\bar{x}$	Magnitude Rating for Colonoscopy, $\bar{x}$
Harms				
Bleeding	1	13[a]	2.67	2.60
Discomfort of ''prep'' for colonoscopy	NA	21	NA	2.17
False negatives	4	5	2.75	3.00
False positives	28	6[a]	2.92	3.20
Financial cost	0	19[a]	2.06	2.40
Follow-up procedures	47	10[a]	3.11	2.67[a]
Impotence	19	NA	3.20	NA
Incontinence	21	NA	3.41	NA
Increased mortality	3	4	4.00	3.69
Overdiagnosis	28	8[a]	3.48	2.67[a]
Pain	13	13	2.71	2.08
Perforation	NA	58	NA	3.48
Psychological effects (e.g., anxiety)	53	21[a]	2.83	2.38
Unnecessary treatment	56	11[a]	3.43	2.64[a]
Benefits				
Early detection/treatment	72	74	3.02	3.81[a]
Knowledge/having more information	8	9	3.11	2.90
Lifesaving/reduced mortality	12	13	3.23	3.33
Longevity	12	21[a]	2.75	3.87[a]
Prevent cancer	3	12[a]	2.20	3.68[a]
Psychological effects (e.g., peace of mind)	37	18[a]	2.81	2.95
Rule out cancer	3	7[a]	2.00	3.13[a]

Note: Clinicians rated the magnitude of benefit [harm] on a 4-point scale ranging from *almost no benefit [harm] to patient* (coded as 1) to *large benefit [harm] to patient* (coded as 4). NA = not applicable.
a. $P < 0.001$.

PSA testing (e.g., missing work, distraction from other important healthy activities). Most clinicians (90%) listed a PSA testing harm from at least 1 category from the taxonomy of screening harms,[1] and 65% listed harms in at least 2 categories. Few clinicians cited PSA testing harms from more than 3 categories of the taxonomy (Table 2).

Clinicians listed more harms than benefits of PSA testing ($\bar{x} = 3.03$ v. 1.57, $P < 0.001$). The magnitude of PSA harms was greater than the magnitude of PSA benefits ($\bar{x} = 8.92$ v. 7.16, $P < 0.001$). Mean PSA testing gist indicated that clinicians listed more harms with greater magnitude than benefits ($\bar{x} = -4.12$, $s = 5.56$). Clinicians estimated that getting a PSA test was more likely to harm men than to lengthen their lives ($\bar{x} = 4.41$ v. 2.70, $P < 0.001$) (Table 3).

Colonoscopy

The most frequently mentioned benefits of colonoscopy were early detection/treatment (74%) and longevity (21%). The most frequently listed harms were perforation (58%), discomfort of preparing for the procedure (21%), and psychological effects (e.g., anxiety) (21%) (Table 1). Most clinicians listed at least 1 physical harm of colonoscopy (95%), but fewer recognized psychological harms (29%), hassle (24%), financial strain (19%), or opportunity costs (4%). Most clinicians (88%) listed a colonoscopy harm from at least 1 category from the taxonomy of screening harms,[1] and 44% listed harms in at least 2 categories. Few clinicians cited colonoscopy harms from more than 3 categories (Table 2).

Clinicians listed more harms than benefits of colonoscopy ($\bar{x} = 2.82$ v. 2.02, $P < 0.001$). The magnitude of colonoscopy benefits was greater than the magnitude of harms ($\bar{x} = 8.06$ v. 4.75, $P < 0.001$). Mean colonoscopy gist indicated that clinicians listed more benefits with greater magnitude than harms ($\bar{x} = 0.94$, $s = 4.87$). Clinicians estimated that receiving a colonoscopy was more likely to lengthen life than to cause harm ($\bar{x} = 4.27$ v. 2.45, $P < 0.001$) (Table 3).

PSA Testing versus Colonoscopy

As predicted, clinicians perceived higher likelihood of harm ($z = 8.76$, $P < 0.001$) and lower likelihood of life lengthened ($z = -7.22$, $P < 0.001$) for

Table 2 Proportion of Clinicians Who Listed Harms from Screening

	PSA Testing, %	Colonoscopy, %
Taxonomy category[a]		
Physical harm	70	95
Psychological harm	68	29[b]
Financial strain	13	19[c]
Opportunity cost	0	4
Hassle	56	24
All 5 categories	0	0
Any 4 categories	3	5
Any 3 categories	30	16[b]
Any 2 categories	65	44[b]
Any 1 categories	90	88
No harms listed	10	12

a. Categories are based on the Harris taxonomy of harms (Harris and others, 2014).
b. $P < 0.05$.
c. $P \leq 0.001$.

Table 3 Clinicians' Evaluation of PSA and Colonoscopy

	PSA, $\bar{x}$ (*s*)	Colonoscopy, $\bar{x}$ (*s*)
Gist	−4.12 (5.56)	0.94 (4.87)[a]
Number of harms	3.03 (1.52)	2.82 (1.45)
Number of benefits	1.57 (0.72)	2.02 (1.22)[a]
Summed magnitude of harm	8.92 (4.70)	7.16 (3.91)[a]
Summed magnitude of benefit	4.75 (2.65)	8.06 (3.95)[a]
Subjective difficulty of recall	1.42 (0.76)	1.44 (0.74)
Likelihood of harm	4.41 (2.29)	2.45 (1.47)[a]
Likelihood of life lengthened	2.70 (1.74)	4.27 (2.62)[a]

Note: Gist was the summed magnitude of benefits minus the summed magnitude of harms.
a. $P < 0.001$.

PSA testing relative to colonoscopy. Clinicians' gist of screening was more negative for PSA testing relative to colonoscopy ($z = -8.21$, $P < 0.001$). Considering the components of gist, the summed magnitude of harms clinicians listed was greater ($z = 3.90$, $P < 0.001$) and the summed magnitude of benefits lower ($z = -8.80$, $P < 0.001$) for PSA testing relative to colonoscopy. Clinicians listed fewer benefits ($z = -3.78$, $P < 0.001$) for PSA testing compared with colonoscopy. Clinicians did not perceive the number of harms ($z = 1.42$, $P = 0.16$) or the difficulty of recall ($z = -0.32$, $P = 0.90$) to be different between screening tests (Table 3, Figure 1).

Mediation Analyses

We used separate mediation models to test our 2 competing mediation hypotheses (availability v. fuzzy trace theory). Specifically, mediation models tested whether 1) screening test predicted perceptions of likelihood; 2) screening test predicted potential mediators; and 3) potential mediators predicted likelihood estimates controlling for the effect of screening test on likelihood estimates. The above section, ''PSA Testing versus Colonoscopy,'' shows the results for steps 1 and 2 and indicates that gist and number of benefits were potential mediators. Number of harms and difficulty recalling harms were not candidate mediators, because they failed in step 2 (i.e., were not associated with perceptions of likelihood). We ran additional analyses to examine whether magnitude of benefits were mediators for the sake of completeness, although they were only indirectly part of our mediation hypothesis (as components of gist). Results for the third and fourth steps of the mediation analyses are below.

Effects of Potential Mediators on Perceived Likelihood of Harm

The more positive was clinicians' gist of screening, the lower was their perceived likelihood of harm from screening, controlling for the effect of screening test ($z = -1.91$, $P < 0.05$). In a model that controlled for gist, clinicians estimated that more men would be harmed from PSA testing relative to colonoscopy ($z = 7.44$, $P < 0.001$) (Figure 1). The Sobel test indicated that gist mediated the relationship between screening test and perceived likelihood of harm from screening ($z = -0.25$, $P < 0.05$). Furthermore, the Sobel test showed that the number of benefits ($z = 0.17$, $P < 0.05$) and magnitude of benefit ($z = 0.25$, $P < 0.001$) also mediated this relationship. There were no mediation effects of the number of harms, magnitude of harms, or difficulty of recall.

Effects of Potential Mediators on Perceived Likelihood of Life Lengthened

In a model that controlled for gist, clinicians estimated that fewer men would have their life lengthened from PSA testing than colonoscopy ($z = -4.67$, $P < 0.001$) (Figure 1). The Sobel test indicated that gist mediated the relationship between screening test and perceived likelihood of life lengthened ($z = 0.41$, $P < 0.05$). Furthermore, the Sobel test showed that the number of benefits ($z = -0.26$, $P < 0.05$) and magnitude of benefits ($z = -0.37$, $P < 0.05$) also mediated this relationship. Similar to our findings for perceived likelihood of harm, we found no mediation effects of the number of harms, magnitude of harms, or difficulty of recall.

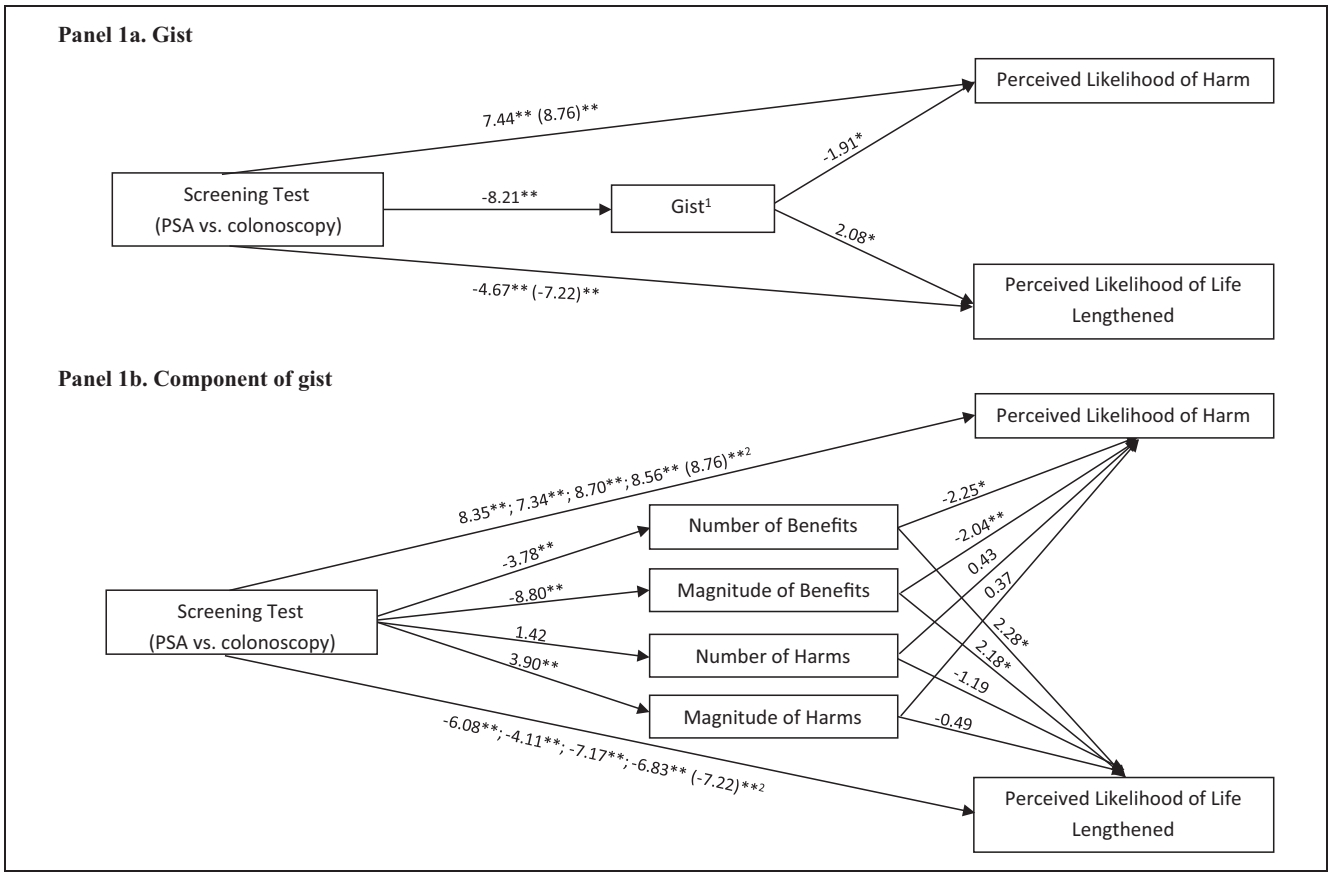

Figure 1 Relationship between screening test and perceived likelihood. Numbers are z *statistics from separate mediation models controlling for survey order. [1]Gist = magnitude of benefits minus magnitude of harms. [2]Effect of screening test on likelihood perceptions controlling for number of benefits; magnitude of benefits; number of harms; magnitude of harms. Main effects in parentheses.* *P $<$ *0.05.* **P $\leq$ *0.001.*

DISCUSSION

Clinicians' perceptions of the likelihood that screening will help or harm play an important role in shaping their screening recommendations. Findings suggest that clinicians were aware of the potential harms of screening but that they had low awareness of the different types of harms. Clinicians in our study judged PSA testing to be more likely to cause harm and less likely to lengthen life relative to colonoscopy, and their gist impressions, mainly of screening benefits, mediated these judgments. Targeting benefits and gist may be the most effective ways to change clinicians' risk perception and screening practices.

Our study is consistent with previous studies showing that clinicians recognize the importance of communicating the harms of cancer screening,[41-43] but our study provides new evidence that clinicians can identify some screening harms with ease. As a group, clinicians listed harms from all categories of the screening cascade identified in the Harris taxonomy.[1] For PSA testing, clinicians listed mostly psychological harms of testing (e.g., anxiety, false positives), physical harms of distal follow-up procedures (e.g., impotence, incontinence), and hassle of unnecessary testing and procedures, suggesting that these types of PSA harms are most available for clinicians. For colonoscopy, clinicians listed mostly physical harms related to the procedure itself (e.g., discomfort of preparation, perforation, bleeding), suggesting that physical harms of colonoscopy are most available for clinicians. However, individual clinicians were less likely to list the full scope of screening harms. Few clinicians listed more than 2 harms of any type. As well, few clinicians enumerated anything beyond physical harms of colonoscopy, and fewer mentioned financial strain or

opportunity costs for either screening test. Clinicians may be unaware of these latter harms or may not think them worth enumerating for themselves or their patients. These findings suggest that if the full spectrum of harms are important to screening decisions, messages to clinicians should emphasize the full scope of harms.[1]

Our study also provides new evidence to increase our understanding of clinicians' perceptions of the benefits of cancer screening. Benefits are important because they play a role in shaping clinicians' screening recommendations,[44,45] which in turn affect patients' screening decisions.[2,3] Our findings suggest that for PSA testing and colonoscopy, clinicians perceived that early detection and treatment and saving lives were the most important benefits, and that early detection and treatment was the most available benefit. This finding is not surprising given that the primary goal of cancer screening is to reduce deaths due to cancer, thereby increasing patients' length of life, as well as curtailing the development of symptomatic metastatic disease.[30,31] However, fewer clinicians listed the psychological benefits of screening (e.g., peace of mind) or enumerated longevity, preventing cancer, ruling out cancer, or having more information. Developing a parallel framework of screening benefits similar to the taxonomy of screening harms developed by Harris and colleagues[1] and testing it to determine patients' values for various benefits could help researchers understand screening benefits, facilitate comparison to screening harms, and ultimately facilitate decision making.

Our study further shows that clinicians relied on the gist they had of screening to formulate their estimations of the likelihood of benefit and harm from screening. These findings offer more support for our fuzzy trace hypothesis than for the availability hypothesis. This mediating role of gist is not surprising given past research showing that physicians frequently rely upon gist when making decisions.[11–14] Increased reliance on gist-based reasoning may reduce errors in probability judgment[46,47] and decrease unhealthy decisions.[48,49] Research has shown that gist is malleable[50,51] and can therefore be targeted for change. Thus, targeting clinicians' gist of screening, for instance through graphical displays that allow clinicians to make gist-based relative magnitude comparisons and detect overarching patterns,[52] could influence their risk perception and possibly reduce overrecommendation of screening.

Interestingly, breaking gist down into its component parts, we found that clinicians' perceptions of screening benefits played a mediating role, but harms did not. This finding deserves further exploration, as it diverges from past research showing that harms information alters risk perception but benefits information does not.[53–55] For clinicians, benefits may be more congruent than harms with thinking about screening tests. That is, clinicians may more easily extract and remember a gist for screening benefits, but they may struggle to extract a gist for screening harms and rely more often on something closer to verbatim memory. The upshot would be that they rely upon benefits (gist) when making likelihood judgments. Indeed, clinicians perceived more abstract, ''gist-like'' benefits (e.g., peace of mind, saving lives) and more concrete harms (e.g., impotence, incontinence, colonic perforation). Alternatively, this finding could reflect that our outcome variable was limited to the likelihood of physical harm but the harms clinicians listed and rated included psychological effects. This finding suggests that messages to clinicians to decrease overrecommendation of screening may need to focus on benefits rather than harms. Future research on the role of gist-based thinking in decision making should assess whether patients and providers formulate risk perceptions in this gist-like way. For example, risk perceptions about certain typically beneficial behaviors such as screening may elicit benefits-based decisions while typically harmful actions like drug use elicit decisions based upon harms but not benefits.

Strengths of our study are the rigorous study design and high response rate. Furthermore, the within-subjects design controlled for individual differences and thus increased statistical power. While the juxtaposition of PSA testing to colonoscopy was informative, these 2 screening tests are qualitatively different in ways that make this comparison an imperfect one. PSA testing and colonoscopy involve different procedures (i.e., blood test v. internal examination) that have different harms that occur at different stages of the screening process. We chose the comparison, however, because we wanted to compare screening tests of varying benefit-to-harm ratio in order to get a more complete understanding of clinicians' likelihood judgments related to screening. It is possible that clinicians may have several gists about screening in addition to the one we calculated in our study. Due to the parent study's focus on harms of clinical preventive services and limited space on the questionnaire, we did not measure clinicians' ease of recalling benefits, preventing us from comparing clinicians' ease of recalling harms to benefits. Findings are from an academic group of clinicians who may be better informed than the average

clinician. We measured availability using a paper survey, but other approaches such as using time to completion on computer-based tasks may be more precise. Another limitation was the relatively high number of missing values for some mediator variables. We acknowledge that results from our hypothetical vignettes may differ from clinicians' reactions to real-life patients that occur under time pressure and complex circumstances. We also acknowledge the need to replicate findings with vignettes that vary the characteristics of hypothetical patients and assess additional screening tests.

Future research should establish whether the findings that we reported here generalize to other populations of clinicians and to beliefs about cancer screening tests delivered solely to women. Future research should also assess clinicians' perceptions of the likelihood of psychological harm from screening, as well as patients' perceptions of these important issues. A final area for future research is to examine how clinicians arrive at cancer screening recommendations.

Nonetheless, our findings are promising in that they suggest that clinicians are aware that cancer screening has potential harms as well as benefits, while still being likely to benefit from information to expand their understanding of the different types of screening harms. Findings may be of special interest to health professionals and health services researchers, who may be interested in potential ways to affect clinicians' risk perception and screening practices. Messages to clinicians to decrease overrecommendation and overuse of PSA testing may need to target clinicians' gist of screening and focus on the test's few benefits (e.g., it does not reduce mortality); our findings suggest that this message is more likely to influence clinicians' behavior than the message that PSA testing has many harms.

ACKNOWLEDGMENT

We thank Jessica Defrank for her helpful feedback on early drafts. We affirm that we have listed everyone who contributed significantly to this work.

REFERENCES

1. Harris RP, Sheridan SL, Lewis CL, et al. The harms of screening: a proposed taxonomy and application to lung cancer screening. JAMA Intern Med. 2014;174(2):281–5.

2. Zarychanski R, Chen Y, Bernstein CN, Hebert PC. Frequency of colorectal cancer screening and the impact of family physicians on screening behaviour. CMAJ. 2007;177(6):593–7.

3. Moran WP, Cohen SJ, Preisser JS, Wofford JL, Shelton BJ, McClatchey MW. Factors influencing use of the prostate-specific antigen screening test in primary care. Am J Manag Care. 2000; 6(3):315–24.

4. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica. 1979;47(2).

5. Weinstein ND. Testing four competing theories of health-protective behavior. Health Psychol. 1993;12(4):324–33.

6. Simon HA. Theories of bounded rationality. In: MacGuire CB, Radnor R, eds. Decision and Organization. Amsterdam: North-Holland Publishing Company; 1972. pp 161–76.

7. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Science. 1974;185(4157):1124–31.

8. Cain DM, Detsky AS. Everyone's a little bit biased (even physicians). JAMA. 2008;299(24):2893–5.

9. Fischhoff B, Slovic P, Lichtenstein S. Lay foibles and expert fables in judgments about risk. In: O'Riordan T, Turner RK, eds. Progress in resource management and environmental planning. Chichester, UK: Wiley; 1981. pp 240–55.

10. McDonald CJ. Medical heuristics: the silent adjudicators of clinical practice. Ann Intern Med. 1996;124(1 Pt 1):56–62.

11. Reyna VF, Chick CF, Corbin JC, Hsia AN. Developmental reversals in risk decision making: intelligence agents show larger decision biases than college students. Psychol Sci. 2014;25(1): 76–84.

12. Reyna VF, Adam MB. Fuzzy-trace theory, risk communication, and product labeling in sexually transmitted diseases. Risk Anal. 2003;23(2):325–42.

13. Reyna VF, Lloyd FJ, Whalen P. Genetic testing and medical decision making. Arch Intern Med. 2001;161(20):2406–8.

14. Reyna VF, Lloyd FJ. Physician decision making and cardiac risk: effects of knowledge, risk perception, risk tolerance, and fuzzy processing. J Exp Psychol Appl. 2006;12(3):179–95.

15. Tversky A, Kahneman D. Availability: a heuristic for judging frequency and probability. Cogn Psychol. 1973;9(5(2)):207–32.

16. McDowell ME, Occhipinti S, Chambers SK. The influence of family history on cognitive heuristics, risk perceptions, and prostate cancer screening behavior. Health Psychol. 2013;32(11): 1158–69.

17. Peay MY, Peay ER. The evaluation of medical symptoms by patients and doctors. J Behav Med. 1998;21(1):57–81.

18. Poses RM, Anthony M. Availability, wishful thinking, and physicians' diagnostic judgments for patients with suspected bacteremia. Med Decis Making. 1991;11(3):159–68.

19. MacLeod C, Campbell L. Memory accessibility and probability judgments: an experimental evaluation of the availability heuristic. J Pers Soc Psychol. 1992;63(6):890–902.

20. MacLeod AK, Williams JM, Bekerian DA. Worry is reasonable: the role of explanations in pessimism about future personal events. J Abnorm Psychol. 1991;100(4):478–86.

21. Higgins ET, King G. Accessibility of social constructs: information-processing consequences of individual and contextual variability. In: Cantor N, Kihlstrom JF, eds. Personality, Cognition and Social Interaction. Hillsdale, NJ: Erlbaum; 1981. pp 1–26.

22. Tulving E, Pearlstone Z. Availability versus accessibility of information in memory for words. J Verbal Learning Verbal Behav. 1966;5:381–91.

23. Adam MB, Reyna VF. Coherence and correspondence criteria for rationality: experts' estimation of risks of sexually transmitted infections. J Behav Decis Making. 2005;18(3):169–86.

24. Lloyd FJ, Reyna VF. A web exercise in evidence-based medicine using cognitive theory. J Gen Intern Med. 2001;16(2):94–9.

25. Reyna VF. A theory of medical decision making and health: fuzzy trace theory. Med Decis Making. 2008;28(6):850–65.

26. Reyna VF, Estrada SM, DeMarinis JA, Myers RM, Stanisz JM, Mills BA. Neurobiological and memory models of risky decision-making in adolescents versus young adults. J Exp Psychol Learn Mem Cogn. 2001;37(5):1125–42.

27. Sawaya GF, Guirguis-Blake J, LeFevre M, Harris R, Petitti D. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. Ann Intern Med. 2007;147(12):871–5.

28. Rex DK, Johnson DA, Anderson JC, Schoenfeld PS, Burke CA, Inadomi JM. American College of Gastroenterology guidelines for colorectal cancer screening 2009 [corrected]. Am J Gastroenterol. 2009;104(3):739–50.

29. Davila RE, Rajan E, Baron TH, et al. ASGE guideline: colorectal cancer screening and surveillance. Gastrointest Endosc. 2006; 63(4):546–57.

30. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2008;149(9): 627–37.

31. U.S. Preventive Services Task Force. U.S. Preventive Services Task Force recommendation. Available from: URL: http://www.uspreventiveservicestaskforce.org/prostatecancerscreening/prostatefinalrs.htm

32. American Cancer Society. American Cancer Society recommendations for early cancer detection. Available from: URL: http://www.cancer.org/cancer/prostatecancer/moreinformation/prostatecancerearlydetection/prostate-cancer-early-detection-acs-recommendations

33. American Urological Association. AUA releases new clinical guideline on prostate cancer screening. Available from: URL: https://www.auanet.org/advnews/press_releases/article.cfm?articleNo=290

34. Edwards P, Cooper R, Roberts I, Frost C. Meta-analysis of randomised trials of monetary incentives and response to mailed questionnaires. J Epidemiol Commun Health. 2005;59(11):987–99.

35. Elstad EA, Sheridan SL, Lee JG, Rini C, Earp JA, Brewer NT. Have screening harms become newsworthy? *News coverage of prostate and colorectal cancer screening since the* 2008 USPSTF recommendation changes. J Behav Med. 2014;37(6):1242–51.

36. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society 1977;39(1):1.

37. Schafer JL. Multiple imputation: A primer. Statistical Methods in Medical Research 1999;8(1):3.

38. Bodner TE. What improves with increased missing data imputations? Struct Equ Modeling. 2008;15:651–75.

39. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 1986;51(6): 1173–82.

40. SAS 9.1. Cary (NC): SAS Institute Inc.; 2004.

41. Bowen DJ, Hannon PA, Harris JR, Martin DP. Prostate cancer screening and informed decision-making: provider and patient perspectives. Prostate Cancer Prostatic Dis. 2011;14(2):155–61.

42. Deobald R, Graham P, Chad J, et al. Colorectal cancer screening practices in Saskatchewan: survey of family physicians. Can Fam Physician. 2013;59(12):558–63.

43. Smith P, Hum S, Kakzanov V, Del Giudice ME, Heisey R. Physicians' attitudes and behaviour toward screening mammography in women 40 to 49 years of age. Can Fam Physician. 2012;58(9):508–13.

44. Lewis CL, Griffith J, Pignone MP, Golin C. Physicians' decisions about continuing or stopping colon cancer screening in the elderly: a qualitative study. J Gen Intern Med. 2009;24(7): 816–21.

45. Henderson S, DeGroff A, Richards TB, et al. A qualitative analysis of lung cancer screening practices by primary care physicians. J Commun Health. 2011;36(6):949–56.

46. Reyna VF. Class inclusion, the conjunction fallacy, and other cognitive illusions. Dev Rev. 1991;11(4):317–36.

47. Reyna VF, Mills BA. Converging evidence supports fuzzy-trace theory's nested sets hypothesis (but not the frequency hypothesis). Behav Brain Sci. 2007;30(3):278–80.

48. Reyna VF, Farley F. Risk and rationality in adolescent decision-making: implications for theory, practice, and public policy. Psychol Sci Public Interest. 2006;7(1):1–44.

49. Reyna VF. Conceptions of memory development, with implications for reasoning and decision making. Ann Child Dev. 1996; 12:87–118.

50. Fraenkel L, Peters E, Charpentier P, et al. Decision tool to improve the quality of care in rheumatoid arthritis. Arthritis Care Res. 2012;64(7):977–85.

51. Wolfe CR, Reyna VF. Assessing semantic coherence and logical fallacies in joint probability estimates. Behav Res Methods. 2010;42(2):373–80.

52. Stone ER, Sieck WR, Bull BE, Yates JF, Parks SC, Rush CJ. Foreground: background salience: explaining the effects of graphical displays on risk avoidance. Organ Behav Hum Decis Process. 2003;90:19–36.

53. Perneger TV, Schiesari L, Cullati S, Charvet-Berard A. Does information about risks and benefits improve the decision-making process in cancer screening—randomized study. Cancer Epidemiol. 2011;35(6):574–9.

54. Perneger TV, Cullati S, Schiesari L, Charvet-Berard A. Impact of information about risks and benefits of cancer screening on intended participation. Eur J Cancer. 2010;46(12):2267–74.

55. Lee DH, Mehta MD. Evaluation of a visual risk communication tool: effects on knowledge and perception of blood transfusion risk. Transfusion. 2003;43(6):779–87.