Original Article

Medical Decision Making

Medical Decision Making
2018, Vol. 38(3) 355–365

DOI: 10.1177/0272989X17729359
journals.sagepub.com/home/mdm
SAGE

Strengths and Gaps in Physicians' Risk Communication: A Scenario Study of the Influence of Numeracy on Cancer Screening Communication

Dafina Petrova, Olga Kostopoulou, Brendan C. Delaney, Edward T. Cokely, and Rocio Garcia-Retamero

Abstract

Objective. Many patients have low numeracy, which impedes their understanding of important information about health (e.g., benefits and harms of screening). We investigated whether physicians adapt their risk communication to accommodate the needs of patients with low numeracy, and how physicians' own numeracy influences their understanding and communication of screening statistics. **Methods.** UK family physicians ($N = 151$) read a description of a patient seeking advice on cancer screening. We manipulated the level of numeracy of the patient (low v. high v. unspecified) and measured physicians' risk communication, recommendation to the patient, understanding of screening statistics, and numeracy. **Results.** Consistent with best practices, family physicians generally preferred to use visual aids rather than numbers when communicating information to a patient with low (v. high) numeracy. A substantial proportion of physicians (44%) offered high quality (i.e., complete and meaningful) risk communication to the patient. This was more often the case for physicians with higher (v. lower) numeracy who were more likely to mention mortality rates, OR = 1.43 [1.10, 1.86], and harms from overdiagnosis, OR = 1.44 [1.05, 1.98]. Physicians with higher numeracy were also more likely to understand that increased detection or survival rates do not demonstrate screening effectiveness, OR = 1.61 [1.26, 2.06]. **Conclusions.** Most physicians know how to appropriately tailor risk communication for patients with low numeracy (i.e., with visual aids). However, physicians who themselves have low numeracy are likely to misunderstand the risks and unintentionally mislead patients by communicating incomplete information. High-quality risk communication and shared decision making can depend critically on factors that improve the risk literacy of physicians.

Keywords
cancer screening, numeracy, risk communication

Date received: August 8, 2016; accepted: July 27, 2017

Cancer screening can save lives but can also carry risks, such as false positive results and the risk of unnecessary treatment.[1] Given the evidence for potential harm across various screening tests for cancer (e.g., breast, prostate, lung, thyroid[2]), it is likely that informed rather than persuasion-based decision making will become the standard for many screening decisions.[2–4] However, several obstacles to informed decision making have been documented. For example, many US and European adults

Mind, Brain, and Behavior Research Center, University of Granada, Granada, Spain (DP, RG); Department of Surgery and Cancer, Division of Surgery, Imperial College London, London, England, UK (OK, BCD); National Institute for Risk & Resilience, and Department of Psychology, University of Oklahoma, Norman, OK, USA (ETC); and Max Planck Institute for Human Development, Germany (ETC, RG). Financial support for this study was provided in part by grants from Ministerio de Economía y Competitividad (Spain) (PSI2011-22954 and and PSI2014-51842-R to RGR and DP), by the U.S. National Science Foundation (SES–1253263), and by the National Institute for Risk & Resilience at the University of Oklahoma. The study received NHS Health Research Authority approval (IRAS project ID: 189158) for data collection in the UK. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

Corresponding Author
Dafina Petrova, Mind, Brain and Behavior Research Center, Universidad de Granada, Campus Universitario de Cartuja s/n, 18071 Granada, Spain. (dafinapetrova@ugr.es)

believe that cancer screening is almost always beneficial. In turn, they tend to grossly overestimate the benefits and are often unaware of possible harms.[5–11] People also often perceive that screening is an obligation to family and society and that foregoing screening is irresponsible.[9,11] Such beliefs are occasionally reinforced by campaigns that use misleading statistics, exaggerate the benefits, and omit serious harms.[12,13] Highly positive attitudes towards cancer screening can stem from the perceived value of saving a life and the effectiveness of some screening programs; however, such attitudes can also lead to misinformed decisions about screening programs with disputed or mixed efficacy.[14–16] For example, the United States Preventative Services Task Force recommends against screening for prostate cancer with PSA tests.[17] However, even when shown clearly presented evidence that this screening causes harm and is on average ineffective, some men fail to understand the evidence and are willing to get screened.[15]

Research shows that physicians' knowledge of screening benefits and harms is also often incomplete, and so physicians may fail to discuss harms with their patients.[6,18–21] A national survey of US physicians showed that most primary care physicians mistakenly interpreted increased detection and improved survival as evidence that screening saves lives.[22] This shows a lack of awareness that survival statistics are distorted by both lead time bias and overdiagnosis bias, and thus do not provide compelling evidence for screening benefits.[22,23]

These circumstances highlight the need for careful communication of the risk of harms and benefits of screening. However, informed decision making about screening can be difficult. First, the statistics involved may be too complex for both patients and physicians to understand.[22] Second, both patients and physicians may need to consider evidence that is inconsistent with traditional screening campaigns and goes against the common belief that early detection should be the default choice.[4,15,16] Despite a growing literature on risk comprehension and risk literacy, there is little direct evidence examining the factors that foster better understanding and communication in this unique context.

One factor that could strongly influence communication and understanding in clinical settings is numeracy of physicians and patients. Statistical numeracy refers to the ability to understand and evaluate numerical expressions of probability and risk. This ability is robustly related to superior decision making and risk literacy, independently from intelligence and education, across medical and other decision contexts.[24,25-31] For example, patients with low numeracy tend to overestimate their risk of cancer and the benefit of medical treatments, and are less capable of using numerical information to inform their perception of risks and benefits.[32-35]

Although research on patient numeracy is abundant, research on the numeracy of health professionals is less extensive. One may expect that physicians' extensive education prepares them to deal with the basic numerical concepts and complex statistics that they will encounter in their practice.[36] Nevertheless, recent research has shown that many physicians lack numeracy and basic risk literacy.[25,37-39] Low physician numeracy is associated with suboptimal recommendations regarding Medicare D plans,[40] inaccurate inferences about screening test results and risks of side effects,[37,41] and reluctance to communicate numerical information to patients.[42] However, the influence of physician numeracy on risk communication approaches with patients is still largely unknown.[42] In theory, high physician numeracy should facilitate understanding of screening statistics and evidence-based recommendations, and thus improve risk communication and shared decision making.[38] We expect that physicians with higher numeracy will be more likely to understand complex screening outcomes and offer better (e.g., more complete) risk communication.

It has also been suggested that physicians adapt their risk communication to the patient's numeracy level.[38,43-47] For example, patients with low numeracy prefer verbal as opposed to numerical information, which they have trouble understanding.[48] Visual aids like icon arrays have gained popularity because they often increase comprehension among patients with low numeracy, as long as patients have basic graph literacy.[49-51] Several instruments exist that can quickly assess patients' health literacy skills and numeracy skills in primary care settings (e.g., see review by Cokely et al.[52]). Previous research suggests that physicians may intuitively use their patients' education level to predict their patients' knowledge about cancer.[53] However, it is not known how information about the patient's numeracy influences physicians' risk communication. We expected that physicians would adapt their communication style to the patient's numeracy in accord with best practices; e.g., when talking to low numeracy patients, they would be more likely to use words and visual aids and less likely to use numbers.

Finally, an official clinical guideline recommending cancer screening may make physicians more likely to recommend screening, regardless of the evidence of its effectiveness. Research suggests that some physicians order screening for their patients even when they do not believe the screening is life saving, because of strong patient demand, fear of lawsuits, or the belief that it represents the standard of practice.[54-56]

In sum, we investigated 3 possible determinants of physicians' risk communication and advice regarding cancer screening: a) their own numeracy, b) their awareness of the patient's level of numeracy, and c) availability of an official clinical guideline recommending screening. We assessed the influence of these factors on 4 outcome measures: 1) risk communication quality, 2) preferred risk communication format, 3) recommendation regarding screening, and 4) understanding of screening statistics.

Method

We conducted an anonymous online survey of family physicians in the UK, a country where screening is offered within national programs.[11] Although family physicians are not directly involved in the delivery of most screening programs, they are the first point of contact should patients have any questions or doubts regarding the benefits and harms from screening (see www.gov.uk/topic/population-screening-programmes). Thus, they may receive inquiries about screening and should be prepared to address them. Potential participants were invited by e-mail to complete a 15-minute survey about "communication of information regarding cancer screening to patients", in return for a £10 Amazon voucher. E-mail addresses were obtained from a database of physicians who had participated in previous studies conducted by the second author and had indicated their willingness to participate again. In October 2015, we emailed a total of 516 currently practicing family physicians. Follow-up emails were sent to non-respondents 1 week after the original e-mail, until the required sample size was reached. We used G*Power to calculate the required sample size. Based on the expected dependent variables related to risk communication (e.g., scores), we planned for a linear regression analysis to detect a small effect size ($f^2 = 0.09$) with an alpha of 0.05, a power of 80%, and 5 predictors (factors in the experimental design are described below). The minimum sample size required to detect the hypothesized main effects was 149. (In the results section, we report the results based on ordinal and Poisson regressions, which provide intuitive measures of effect size [i.e., odds ratios]. The results obtained from the planned linear regressions were highly similar.)

Design, Materials, and Procedure

Participants first completed standard demographic questions (age, gender, year of medical degree, years of experience in family practice, type of family practice [rural, urban, inter-city]). Subsequently, participants read a brief scenario about "Sam" (a gender-neutral name was purposefully chosen), a fictitious 61-year-old patient, who came for advice regarding a screening test for cancer X. No specific cancer was mentioned to avoid the influence of participants' knowledge about existing cancer screening programs. The depicted benefits and risks were designed to be realistic and plausible given the current state of evidence for some cancer screenings (e.g., for breast or prostate cancers, see Figure 1). The scenario and associated questions were developed with the help of an experienced family physician co-author, and were revised after pilot testing with 2 other family physicians.

The scenario employed a 2 (clinical guideline: present v. absent, see wording in the top part of Figure 1) by 2 (effectiveness of screening: effective v. ineffective, see numbers in the middle part of Figure 1) by 3 (patient numeracy: low v. high v. not specified, see wording in the bottom part of Figure 1) between-participants design, resulting in 12 scenario versions. Participants were randomized to 1 of the 12 versions. Depending on the version, participants read information as described in Figure 1. After reading the scenario, participants answered a series of questions measuring risk communication quality, risk communication format, recommendation, and understanding of cancer screening statistics.

Risk communication quality. Participants were shown the following list of topics and asked to indicate which ones they would definitely discuss with the patient: 1) detection of cancer X with screening, 2) detection of cancer X without screening, 3) mortality from cancer X with screening, 4) mortality from cancer X without screening, 5) false-positive rate with screening, 6) false-positive rate without screening, 7) number of people overdiagnosed and treated unnecessarily with screening, 8) number of people overdiagnosed and treated unnecessarily without screening, 9) none of the above, and 10) another topic.

Based on the topics that they selected, we created a quality index based on guidelines for completeness and interpretability of risk communication.[46,57] Because in this context more risk information does not necessarily reflect better communication, the index considers not only how many pieces of information physicians select but also what information they select to communicate. For instance, balanced discussion of mortality statistics and possible harms from screening characterizes good risk communication, while discussion only of detection rates that omits harms characterizes poor risk communication.

Clinical guideline

Present (half of the sample read this):
Currently, the National Health Service (NHS) recommends 3-yearly screening for cancer X for individuals over 50.

Absent (the other half of the sample read this):
Currently, there is no National Health Service (NHS) guideline regarding screening for cancer X.

Screening effectiveness

"Please, assume that the screening is free at the point of care and to the provider. It is non-invasive, and detects cancer X, for which treatment, such as surgery or chemotherapy, is available. The table below presents data about patients aged 50 to 70. The data come from a large trial of UK adults that lasted about 10 years."

Effective (half of the sample read this):

	With screening	Without screening
Number of people in whom cancer X was detected.	50 per 1000	30 per 1000
Mortality from cancer X.	3 per 1000	5 per 1000
False positive screening tests that were followed by unnecessary biopsies.	100 per 1000	0 per 1000
People overdiagnosed and treated unnecessarily.	18 per 1000	0 per 1000

Not effective (the other half of the sample read this):

	With screening	Without screening
Number of people in whom cancer X was detected.	50 per 1000	30 per 1000
Mortality from cancer X.	5 per 1000	5 per 1000
False positive screening tests that were followed by unnecessary biopsies.	100 per 1000	0 per 1000
People overdiagnosed and treated unnecessarily.	20 per 1000	0 per 1000

Patient numeracy

High (1/3 of the sample read this):
"As part of a new initiative by the NHS, Sam took a short online test before making an appointment. This test showed that Sam has a very good understanding of numbers and probabilities."

Low (1/3 of the sample read this):
"As part of a new initiative by the NHS, Sam took a short online test before making an appointment. This test showed that Sam has a very low understanding of numbers and probabilities."

Unspecified:
No relevant information was given to another 1/3 of the sample.

Figure 1 Information provided to participants based on scenario version. Note: The statistics shown to communicate screening effectiveness were fictitious but based on outcomes of some common cancer screenings (e.g., screening for prostate cancer with PSA tests and screening for breast cancer with mammography[69,70]).

Regarding benefits (total of 2 points), if physicians chose to discuss both detection and mortality with and without screening (1 to 4), they received 2 points. If they chose to discuss mortality with and without screening (3 and 4), without discussing detection (1 and 2), they still received 2 points. The reason is that mortality rates show the best evidence to evaluate screening benefit (i.e., they are unaffected by lead-time and overdiagnosis biases).[22] If physicians chose to discuss detection (1 and 2) but not mortality (3 and 4), they received 0 points, because detection rates by themselves do not convey whether screening saves lives and are influenced by the overdiagnosis bias.[22] If physicians chose to discuss mortality with or without screening only (only 3 or only 4), they received 0 points because such communication is unbalanced and difficult to interpret (e.g., if only mortality rates with screening are discussed, it is not clear if these are smaller than without screening).[46]

Regarding harms (total of 2 points), if physicians chose to discuss false-positive rates with screening (5), they received 1 point. If they chose to discuss numbers of people overdiagnosed with screening (7), they also received 1 point. Few physicians chose to discuss none of the suggested topics (9) or another topic (10), so we did not consider these further.

This resulted in a final score that gave equal weight to benefits (2 points max.) and harms (2 points max.) and ranged from 0 (worst risk communication) to 4 (best risk communication).

Risk communication format. For each topic selected above, participants were asked to indicate their preferred format of risk communication. They were presented with the following options: conveying the information (a) with words (e.g., small, larger, etc.), (b) with visual aids (e.g., draw a simple graph, use an icon array), and/or (c) with numbers (e.g., percentages, number of people). Participants could choose more than one communication format. We counted how many times they chose each format.

Recommendation. Participants indicated their screening recommendation to the patient by choosing either "in favor", "neutral" (neither for nor against), or "against".

Table 1 Items Used to Assess Physician Understanding of Cancer Screening Statistics from Wegwarth et al[22,a]

	What proves that a cancer screening test "saves lives"?		
	Proves	**Does not prove**	**I don't know**
More cancers are detected in screened populations than in unscreened populations.	24 (16)	121 (80)[b]	6 (4)
Cancers detected at screening have better 5-year survival rates than cancers detected because of symptoms.	77 (51)	64 (42)[b]	10 (7)
In a randomized trial, mortality rates of screened persons are lower than those of unscreened persons.	100 (66)[b]	36 (24)	15 (10)

[a]Numbers of Physicians (% of the total sample; $N = 151$).
[b]The correct answer to each question.

Understanding of cancer screening statistics. This was measured using a short questionnaire.[22] For each of the 3 statements presented in Table 1, physicians indicated whether it proved or not that screening saves lives. We computed the sum of correct answers (range, 0 to 3).

Physician numeracy. This was measured with the Berlin Numeracy Test-Schwartz (BNT-S).[25] The test consists of 2 to 3 relatively difficult items from the adaptive BNT (see RiskLiteracy.org) and 3 relatively easy items from Schwartz et al.[32] The BNT is suited for use in highly educated samples and the 3 easy items (administered first) allow for higher discriminability among individuals with low-to-moderate numeracy. The final score ranged from 1 to 7, where a higher score indicated higher numeracy.

Analysis

We conducted multiple regressions in SPSS 21. The independent variables in all regressions were physician numeracy (continuous variable), presence of a clinical guideline (present v. absent), effectiveness of screening (effective v. ineffective), level of patient numeracy (low v. high v. unspecified), and years of experience as a family physician (control variable). The dependent variables were risk communication quality (a); risk communication format – number of times physicians chose each format: words (b), visual aids (c), and numbers (d); recommendation (e); and understanding of screening statistics (f), resulting in 6 multiple regression analyses. We present the essential results below. More detailed regression results are included in Appendix Tables S2 and S3.

As a measure of effect size, we computed odds ratios (OR) and 95% confidence intervals (CI). For physician numeracy, we computed ORs representing 1 unit change on the 1 to 7 scale of measurement. We also computed ORs that represented a 6-unit change on the 1 to 7 scale to provide a more intuitive measure of effect size, comparing the odds of physicians with the highest v. low numeracy: $OR_{\text{min-max}}$. For multiple group comparisons, we used Bonferroni-corrected alpha.

Results

Of 516 invited family physicians, 174 (34%) started the survey and 151 (87%) completed it, reaching the planned sample size within 3 weeks. Half of the participants (50%) were female and most (52%) worked in urban practices. The sample included physicians of various ages: 25 to 34 y (13%), 35 to 44 y (54%), 45 to 54 y (20%), and 55 y and over (13%). The gender and age distribution of the sample was similar to nation-wide statistics for family physicians (see www.gmc-uk.org), except for the 35 to 44 y age group, which was overrepresented. Complete characteristics of the sample and descriptive statistics for the outcome measures are found in Tables 1, 2, and Appendix Table S1. Physician numeracy was high on average (see Table 2), with 40% of physicians scoring the maximum of 7 on the BNT-S, 43% scoring 5 or 6, and only 16% scoring 4 or below. Risk communication quality was also satisfactory on average, with 44% of physicians offering risk communication of high quality (score = 4), 36% of moderate quality (score = 2 or 3), and 20% of low quality (score = 0 or 1).

Risk Communication Quality

We conducted an ordinal regression analysis with the risk communication quality index as a dependent variable (range, 0 to 4; see also Appendix Table S2). Physician numeracy was the only factor predicting risk communication quality, such that physicians with higher numeracy

Table 2 Key Variables in the Dataset (N = 151)

	Minimum	Maximum	Mean	SD
Physician age	30	65	42.3	8.2
Physician experience: number of years in family practice	2	35	12.2	8.3
Physician numeracy (the Berlin Numeracy Test-Schwartz)	1	7	5.7	1.3
Risk communication quality	0	4	2.8	1.3
Total number of risk communication topics selected	1	8	4.8	1.9
Risk communication format: choosing words	0	8	1.3	1.9
Risk communication format: choosing visual aids	0	8	1.2	2.0
Risk communication format: choosing numbers	0	8	2.3	2.5
Understanding of screening statistics score	0	3	1.9	0.8

SD, standard deviation.

offered better risk communication: OR = 1.39 [1.10, 1.76], $OR_{min\text{-}max}$ = 7.21 [1.77, 29.72]; P = 0.007.

We performed follow-up logistic regressions on each of the topics that physicians could choose to discuss with the patient (i.e., detection, mortality, false positives, and overdiagnosis). For this analysis, if risks both with and without screening were mentioned, detection and mortality were coded as discussed (score of 1). If they were not mentioned at all or only risks with or without screening were mentioned, detection and mortality were coded as not discussed (score of 0). False positives and overdiagnosis were coded as discussed (score of 1) if the risks with screening were mentioned and as not discussed (score of 0) if this communication was omitted. Higher physician numeracy was associated with higher odds of discussing (v. not) mortality rates (OR = 1.43 [1.10, 1.86], $OR_{min\text{-}max}$ = 8.55 [1.77, 41.41], P = 0.007), and harms from overdiagnosis (OR = 1.44 [1.05, 1.98], $OR_{min\text{-}max}$ = 8.92 [1.34, 60.25], P = 0.023). Numeracy was not related to the discussion of detection rates (OR = 1.19 [0.93, 1.52], P = 0.178) or false positives (OR = 1.13 [0.85, 1.50], P = 0.417).

Risk Communication Format

We performed 3 Poisson regressions, with the number of times physicians chose to communicate with words, visual aids, and numbers, respectively. In these regressions, we also controlled for the number of topics in total that physicians chose to discuss (see Appendix Table S3).

Words. Patient numeracy had no significant effect on how often physicians chose words. Physicians chose to communicate with words more often when there was an official guideline recommending the screening (OR = 1.51 [1.12, 2.03]; P = 0.006) and when the screening was not life-saving (OR = 1.48 [1.10, 1.99], P = 0.009). There were no other significant effects ($P > 0.05$).

Visual aids. If the patient was described as having low (v. high) numeracy, physicians more often chose to communicate risk using visual aids (OR = 2.78 [1.89, 4.08], $P < 0.001$) (see Table 3). Increasing physician experience was related to choosing visual aids more often (OR = 1.03 [1.01, 1.05], $P < 0.001$; OR for 10 years of experience was 1.34 [1.10, 1.63]). There were no further significant effects ($P > 0.05$).

Numbers. If the patient was described as having low (v. high) numeracy, physicians chose to use numbers less often (OR = 0.59 [0.45, 0.77], $P < 0.001$; see Table 3). There were no further significant effects ($P > 0.05$).

Recommendation

We performed an ordinal regression with recommendation as the dependent variable (not recommend v. neutral v. recommend, see Appendix Table S2). Thirty-five (23%) physicians recommended screening to the patient, 57 (38%) recommended not to get screened, and 59 (39%) were neutral. Physicians tended to recommend screening more often when there was a clinical guideline (27%) than when there was not (19%) but this relationship was not significant: OR = 1.68 [0.91, −3.11], P = 0.096. Twenty percent of physicians recommended screening when it was not effective (v. 27% who recommended it when it was effective, P = 0.061). There were no further significant effects ($P > 0.05$).

Understanding of Screening Statistics

We performed ordinal regression with understanding of screening statistics as a dependent variable (range 0 to 3,

Table 3 Estimated Marginal Means for How Many Times Physicians Chose to Communicate the Screening Outcomes They Selected for Discussion: Words, Visual Aids or Numbers as a Function of the Patient's Numeracy Level[a]

			Comparisons of Patient Numeracy Levels					
			Low v. High		Low v. Unspecified		Unspecified v. High	
	Patient Numeracy	Mean (SE)	OR [LLCI, ULCI]	*P*	OR [LLCI, ULCI]	*P*	OR [LLCI, ULCI]	*P*
Words	Low	1.10 (0.14)	1.02 [0.71, 1.46]	0.907	0.85 [60, 1.21]	0.364	1.20 [0.85, 1.70]	0.309
	Unspecified	1.29 (0.16)						
	High	1.07 (0.14)						
Visual aids	Low	1.72 (0.18)	2.78 [1.89, 4.08]	< 0.001*	1.79 [1.28, 2.52]	0.001	1.55 [1.02, 2.36]	0.040
	Unspecified	0.96 (0.14)						
	High	0.62 (0.11)						
Numbers	Low	1.55 (0.17)	0.59 [0.45, 0.77]	< 0.001*	0.77 [0.58, 1.02]	0.064	0.77 [0.60, 0.98]	0.035
	Unspecified	2.03 (0.20)						
	High	2.64 (0.22)						

LLCI/ULCI, lower level/upper level 95% CI; SE, standard error of the mean.
[a]Odds ratios (OR) are from Poisson regressions and compare the number of times physicians chose each format between the different patient numeracy conditions. Significant comparisons are marked with asterisk according to Bonferroni-corrected alpha level 0.016.

see Appendix Table S2). Physician numeracy was the only factor predicting understanding of screening statistics: OR = 1.61 [1.26, 2.06], $OR_{\text{min-max}}$ = 17.41 [4.01, 76.42], $P < 0.001$.

We performed follow-up logistic regressions on the 3 separate items of the test. With increasing numeracy, physicians were more likely to know that neither increased detection (item 1, OR = 1.71 [1.25, 2.34], $OR_{\text{min-max}}$ = 25.00 [3.81, 164.17], $P = 0.001$) nor survival rates (item 2, OR = 1.35 [1.04, 1.75], $OR_{\text{min-max}}$ = 6.05 [1.27, 28.72], $P = 0.026$) demonstrated screening effectiveness. Numeracy was not associated with knowing that decreased mortality found in a randomized trial demonstrated screening effectiveness (item 3, OR = 1.21 [0.93, 1.55], $P = 0.151$).

We tested a mediation mechanism such that physicians' understanding of screening statistics explained the effect of physician numeracy on risk communication quality. Higher physician numeracy was related to better knowledge of screening statistics ($\beta = 0.33$, $P < 0.001$), and better knowledge of screening statistics was related to higher quality risk communication ($\beta = 0.19$, $P = 0.021$). However, once physician numeracy was added as a predictor of risk communication ($\beta = 0.20$, $P = 0.019$), knowledge of screening statistics was no longer significantly related to risk communication ($\beta = 0.12$, $P = 0.148$). These results suggest that the relationship between knowledge and risk communication was explained by physician numeracy, which had an independent positive effect on both outcomes.

Discussion

This study demonstrated how numeracy shapes physicians' risk communication. Most physicians offered high or moderate quality risk communication to the hypothetical patient. Yet, a substantial minority chose to communicate incomplete and possibly misleading information. Importantly, this minority was characterized by low levels of numeracy.

Physicians with the highest (v. lowest) numeracy in our sample had 7 times the odds of offering complete and balanced information about screening to patients – an effect size indicating high clinical relevance. Specifically, they were more likely to communicate mortality rates and risks from overtreatment. Physicians with higher numeracy were also less likely to mistakenly believe that increased detection or survival rates from screening show that screening saves lives.[22] These results are in line with research showing that physician numeracy is related to better diagnostic inferences about the predictive value of screening tests and better comprehension of important surgical risks.[37,41]

Overall, these results clearly show the benefits of physician numeracy for fostering risk literacy and evidence-based decisions in clinical practice. Results also suggest that physicians with low numeracy may generally be ill-equipped to facilitate informed decision making about cancer screening and other medical treatments that require the consideration of numerical evidence. A previous study linked physicians' lower data interpretation abilities (e.g., distinguishing between relative and absolute risk) to more enthusiasm about cancer screening in general.[58] Although we did not find a relationship between physician numeracy and screening recommendations, our findings suggest that physicians with low numeracy are much more likely to provide patients with insufficient and lower quality information about screening.

It is encouraging that physicians successfully adapted their risk communication to suit the patient's level of numeracy, and this effect was largely independent of their own numeracy level. There is evidence that visual aids often enhance the understanding of risk by patients with low numeracy.[49] Our participants' choice of visual aids as the most preferred mode of communicating risk to low numeracy patients suggests that physicians should have the necessary tools to assess patient numeracy where appropriate, and are both trained and willing to use visual aids for communicating screening benefits and harms. Even if physicians do not know that visual aids make risks easier to comprehend, these powerful and simple tools would be a ready risk communication means at their disposal.

The presence of an official guideline was related to a small increase in recommendations in favor of screening. There was a similar small increase in recommendations when screening was life-saving v. not. These 2 independent effects are in line with previous findings showing that physicians sometimes order screening for their patients because they believe this is the standard of practice, even if they are not convinced that screening improves patient outcomes.[56] Basic psychological mechanisms potentially driving these effects are the persuasiveness of expert sources[59] and its relation to the thoughtful consideration of the evidence.[60] For instance, physicians may have recommended screening in the ineffective condition because they did not sufficiently process the evidence in the hypothetical scenario, leading to simple reliance on the guideline. Alternatively, physicians may have decided to follow the guideline, despite sufficiently processing the information given (e.g., "I must be misunderstanding" or "This is what I am supposed to do"). The latter may also be related to defensive decision making practices of physicians. For instance, physicians tend to select much more conservative treatments for their patients than for themselves for fear of legal consequences.[24]

Limitations

Scenarios like the one used here tend to be a valid and reliable assessment of clinical practice, being a useful proxy for actual behavior.[61] However, there are many relevant factors that are likely to influence screening communication and recommendations that were not considered in the current study. These include actual and perceived patient demand, the availability of sufficient time for discussion with the patient, and access to communication resources, including relevant data or visual aids. The influence of physician numeracy should be further explored in richer scenarios or actual information exchanges. To the best of our knowledge, no research has investigated the role of physician numeracy in actual risk communication to patients. It is also not clear how often risk communication about screening takes place in practice. However, the demand for such communication is likely to increase given the growing emphasis on informed decision making and the increasing evidence of harms from some screenings. The current study suggests that physicians with low numeracy need assistance to meet such demands properly.

Although the sample represented physicians with different demographic backgrounds and years of experience, it is possible that physicians who have low numeracy are underrepresented due to low willingness to participate in a study that involves numerical information. To reduce this possibility, our study invitation contained no mention of numerical information. Whereas only 34% of physicians who were invited started the survey, as much as 87% of those who did completed it. Only 8 out of 174 (5%) participants left the study on the information page where the numeracy assessment was mentioned. Nevertheless, physicians may have expected to see numerical information in the context of cancer screening, which may have discouraged less numerate physicians from participating. This implies that the documented deficits related to numeracy may be even larger in the broader sampling population.

The index of quality of risk communication was calculated based on published guidelines for complete and transparent risk communication.[46,57] These guidelines were applied to the current context based on our best expert judgment. Nevertheless, the resulting index may not be directly applicable to all clinical situations and alternative formulations are possible (e.g., giving more weight to benefits v. harms or vice versa as a function of the clinical context).

As with every research, it is possible that instead of giving their true opinion, participants may have tried to mark the "correct" response sought by the experimenter. We find this to be unlikely for several reasons. First, if that were the case, participants would have tended to select all suggested topics for discussion with the patient.

Instead, participants selected on average 5 out of 8 topics (Table 2). Second, the survey was anonymous, and participants were assured that their responses could not be linked to their email addresses.

Implications

Numeracy is an essential component of risk literacy that tends to facilitate informed, evidence-based decisions.[25,62] Unfortunately, not all patients and physicians have developed the skills that are necessary for independently evaluating and understanding the complex statistics that are now part of many medical decisions. On one hand, the current research suggests that screening patients for numeracy may help many physicians tailor risk communication to patient needs and abilities, partially mitigating this problem.[57,63] On the other hand, this research shows how insufficient physician numeracy can impede patients' informed decision making across various contexts, because, as is common among many diverse professionals, many physicians are simply unaware of their misunderstanding.[29] Put simply, although physicians with low numeracy know how to make risks easier to understand for patients, they themselves have trouble understanding, and can thus unintentionally mislead their patients by communicating incomplete information. A telling example is the worrying 20% of physicians in our study who recommended that a hypothetical patient participate in screening that was not effective and that was likely to cause serious harm. To avoid such misguided recommendations, it is necessary to improve numeracy, risk literacy, and statistical skills training in medical curricula and continuing education.[62,64,65]

A recent review of cancer prevention and screening recommendation statements showed that as much as 69% either did not quantify the benefits and harms of screening or presented them in an asymmetric manner.[66] This, together with the current results, suggests that there is a pressing need for well-designed decision aids to help less numerate physicians and patients understand and discuss life-altering risks and benefits. Examples are simple facts boxes that display the most important information in a tabular format, accompanied by visual aids in the form of icon arrays[14,67,68] (see Harding Center for Risk Literacy—www.harding-center.mpg.de—for examples of good risk communication materials for several cancer screenings). Such decision aids can be easily implemented in clinical practice and can effectively increase comprehension among both patients and physicians.[15,37] Comprehension, in turn, can help patients and physicians discuss and evaluate risks and benefits, in light of the patient's values and informed by the physician's expertise, laying the grounds for shared decision making.

Acknowledgments

We thank David Contreras, Frankie Berry, Victoria Willis, and Renata Samulnik for administrative support.

Supplementary Material

Supplementary material for this article is available on the *Medical Decision Making* Web site at http://journals.sagepub.com/home/mdm.

References

1. Harris RP, Sheridan SL, Lewis CL, et al. The harms of screening: A proposed taxonomy and application to lung cancer screening. *JAMA Int Med*. 2014;174:281–6.
2. Esserman LJ, Thompson IM, Reid B. Overdiagnosis and overtreatment in cancer: An opportunity for improvement. *JAMA*. 2013;310:797–8.
3. Sheridan SL, Harris RP, Woolf SH. Shared decision making about screening and chemoprevention: A suggested approach from the US preventive services task force. *Am J Prev Med*. 2004;26:56–66.
4. Woloshin S, Schwartz LM, Black WC, Kramer BS. Cancer screening campaigns — getting past uninformative persuasion. *N Engl J Med*. 2012;367:1677–9.
5. Gigerenzer G, Mata J, Frank R. Public knowledge of benefits of breast and prostate cancer screening in Europe. *J Nat Cancer Inst*. 2009;101:1216–20.
6. Hoffman RM, Lewis CL, Pignone MP, et al. Decision-making processes for breast, colorectal, and prostate cancer screening: The DECISIONS survey. *Med Decis Making*. 2010;30:53S–64S.
7. Hudson B, Zarifeh A, Young L, Wells JE. Patients' expectations of screening and preventive treatments. *Ann Fam Med*. 2012;10:495–502.
8. Schwartz LM, Woloshin S, Sox HC, Fischhoff B, Welch HG. US women's attitudes to false positive mammography results and detection of ductal carcinoma in situ: Cross sectional survey. *BMJ*. 2000;320:1635–40.
9. Schwartz LM, Woloshin S, Fowler Jr FJ, Welch HG. Enthusiasm for cancer screening in the United States. *JAMA*. 2004;291:71–8.
10. Hersch J, Jansen J, Barratt A, et al. Women's views on overdiagnosis in breast cancer screening: A qualitative study. *BMJ*. 2013;346:f158.
11. Waller J, Osborne K, Wardle J. Enthusiasm for cancer screening in Great Britain: A general population survey. *Br J Cancer*. 2015;112:562–6.

12. Woloshin S, Schwartz LM. How a charity oversells mammography. *BMJ*. 2012;345:e5132.
13. Gigerenzer G. Breast cancer screening pamphlets mislead women. *BMJ*. 2014;348:g2636.
14. Arkes HR, Gaissmaier W. Psychological research and the prostate-cancer screening controversy. *Psychol Sci*. 2012;23:547–53.
15. Petrova D, Garcia-Retamero R, Cokely ET. Understanding the harms and benefits of cancer screening: A model of factors that shape informed decision making. *Med Decis Making*. 2015;35:847–58.
16. Petrova D, Garcia-Retamero R, Catena A, van der Pligt J. To screen or not to screen: What factors influence complex screening decisions? *J Exp Psychol Appl*. 2016;22:247–60.
17. Moyer VA. Screening for prostate cancer: US preventive services task force recommendation statement. *Ann Intern Med*. 2012;157:120–34.
18. Wegwarth O, Gigerenzer G. Overdiagnosis and overtreatment: Evaluation of what physicians tell their patients about screening harms. *JAMA Int Med*. 2013;173:2086–7.
19. Wegwarth O, Gigerenzer G. "There is nothing to worry about": Gynecologists' counseling on mammography. *Patient Educ Couns*. 2011;84:251–6.
20. Han PKJ, Kobrin S, Breen N, et al. National evidence on the use of shared decision making in prostate-specific antigen screening. *Ann Fam Med*. 2013;11:306–14.
21. Elstad EA, Sutkowi-Hemstreet A, Sheridan SL, et al. Clinicians' perceptions of the benefits and harms of prostate and colorectal cancer screening. *Med Decis Making*. 2015;35: 467–76.
22. Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Ann Int Med*. 2012;156:340–9.
23. Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? *JAMA*. 2000;283:2975–8.
24. Garcia-Retamero R, Galesic M. Transparent Communication of Health Risks: Overcoming Cultural Differences. New York: Springer; 2013.
25. Cokely ET, Galesic M, Schulz E, Ghazal S, Garcia-Retamero R. Measuring risk literacy: The Berlin Numeracy Test. *Judg Decis Making*. 2012;7:25–47.
26. Reyna VF, Nelson WL, Han PK, Dieckmann NF. How numeracy influences risk comprehension and medical decision making. *Psychol Bull*. 2009;135:943–73.
27. Garcia-Retamero R, Andrade A, Sharit J, Ruiz JG. Is patients' numeracy related to physical and mental health? *Med Decis Making*. 2015;35:501–11.
28. Cokely ET, Kelley CM. Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judg Decis Making*. 2009;4:20–33.
29. Ghazal S, Cokely ET, Garcia-Retamero R. Predicting biases in very highly educated samples: Numeracy and metacognition. *Judg Decis Making*. 2014;9:15–34.
30. Liberali JM, Reyna VF, Furlan S, Stein LM, Pardo ST. Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *J Behav Decis Making*. 2012;25:361–81.
31. Finucane ML, Gullion CM. Developing a tool for measuring the decision-making competence of older adults. *Psychol Aging*. 2010;25:271.
32. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. *Ann Int Med*. 1997;127:966–72.
33. Davids SL, Schapira MM, McAuliffe TL, Nattinger AB. Predictors of pessimistic breast cancer risk perceptions in a primary care population. *JJ Gen Int Med*. 2004;19:310–5.
34. Lipkus IM, Peters E, Kimmick G, Liotcheva V, Marcom P. Breast cancer patients' treatment expectations after exposure to the decision aid program adjuvant online: The influence of numeracy. *Med Decis Making*. 2010;30:464–73.
35. Woloshin S, Schwartz LM, Black WC, Welch HG. Women's perceptions of breast cancer risk: How you ask matters. *Med Decis Making*. 1999;19:221–9.
36. Anderson BL, Schulkin J. Numerical Reasoning in Judgments and Decision Making about Health. Cambridge: Cambridge University Press; 2014.
37. Garcia-Retamero R, Cokely ET, Wicki B, Joeris A. Improving risk literacy in surgeons. *Patient Educ Couns*. 2016;99:1156–61.
38. Garcia-Retamero R, Wicki B, Cokely ET, Hanson B. Factors predicting surgeons' preferred and actual roles in interactions with their patients. *Health Psychol*. 2014;33:920–8.
39. Anderson BL, Gigerenzer G, Parker S, Schulkin J. Statistical literacy in obstetricians and gynecologists. *J Healthcare Qual*. 2014;36:5–17.
40. Hanoch Y, Miron-Shatz T, Cole H, Himmelstein M, Federman AD. Choice, numeracy, and physicians-in-training performance: The case of Medicare part D. *Health Psychol*. 2010;29:454–9.
41. Garcia-Retamero R, Hoffrage U. Visual representation of statistical information improves diagnostic inferences in doctors and their patients. *Soc Sci Med*. 2013;83:27–33.
42. Anderson BL, Obrecht NA, Chapman GB, Driscoll DA, Schulkin J. Physicians' communication of down syndrome screening test results: The influence of physician numeracy. *Genetics Med*. 2011;13:744–9.
43. Elwyn G, Edwards A, Kinnersley P, Grol R. Shared decision making and the concept of equipoise: The competences of involving patients in healthcare choices. *Br J Gen Pract*. 2000;50:892–9.
44. Galesic M, Garcia-Retamero R. Do low-numeracy people avoid shared decision making? *Health Psychol*. 2011;30: 336–41.
45. Hanoch Y, Miron-Shatz T, Rolison JJ, Omer Z, Ozanne E. Shared decision making in patients at risk of cancer: The role of domain and numeracy. *Health Expect*. 2015;18: 2799–810.
46. Gaissmaier W, Anderson BL, Schulkin J. How do physicians provide statistical information about antidepressants

to hypothetical patients? *Med Decis Making*. 2014;34: 206–15.
47. Schwartz PH. Decision aids, prevention, and the ethics of disclosure. *Hastings Cent. Rep*. 2011;41:30–9.
48. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring numeracy without a math test: Development of the subjective numeracy scale. *Med Decis Making*. 2007;27:672–80.
49. Garcia-Retamero R, Cokely ET. Communicating health risks with visual aids. *Curr Dir Psychol Sci*. 2013;22:392–9.
50. Gaissmaier W, Wegwarth O, Skopec D, Müller A, Broschinski S, Politi MC. Numbers can be worth a thousand pictures: Individual differences in understanding graphical and numerical representations of health-related information. *Health Psychol*. 2012;31:286–96.
51. Okan Y, Garcia-Retamero R, Cokely ET, Maldonado A. Improving risk understanding across ability levels: Encouraging active processing with dynamic icon arrays. *J Exp Psychol Appl*. 2015;21:178–94.
52. Cokely ET, Ghazal S, Garcia-Retamero R. Measuring numeracy. In: Anderson BL, Schulkin J eds. *Numerical Reasoning in Judgments and Decision Making about Health*. Cambridge: Cambridge University Press; 2014.
53. Anderson BL, Schulkin J. Physicians' perceptions of patients' knowledge and opinions regarding breast cancer: Associations with patient education and physician numeracy. *Breast Care*. (Basel) 2011;6:285–8.
54. Austin OJ, Valente S, Hasse LA, Kues JR. Determinants of prostate-specific antigen test use in prostate cancer screening by primary care physicians. *Arch Fam Med*. 1997;6:453–8.
55. Hicks RJ, Hamm RM, Bemben DA. Prostate cancer screening. what family physicians believe is best. *Arch Fam Med*. 1995;4:317–22.
56. Voss JD, Schectman JM. Prostate cancer screening practices and beliefs. *J Gen Int Med*. 2001;16:831–7.
57. Trevena LJ, Zikmund-Fisher BJ, Edwards A, et al. Presenting quantitative information about decision outcomes: A risk communication primer for patient decision aid developers. *BMC Med Inf Decis Making*. 2013;13:S7.
58. Caverly TJ, Prochazka AV, Binswanger IA, Kutner JS, Matlock DD. Confusing relative risk with absolute risk is associated with more enthusiastic beliefs about the value of cancer screening. *Med Decis Making*. 2014;34:686–92.
59. Cialdini RB, Goldstein NJ. Social influence: Compliance and conformity. *Annu Rev Psychol*. 2004;55:591–621.
60. Petty RE, Briñol P. The elaboration likelihood model. *Handbook of Theories of Social Psychology* 2011;1:224–45.
61. Evans SC, Roberts MC, Keeley JW, et al. Vignette methodologies for studying clinicians' decision-making: Validity, utility, and application in ICD-11 field studies. *Int J Clin Health Psychol*. 2015;15:160–70.
62. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients make sense of health statistics. *Psychol Sci Public Int*. 2007;8:53–96.
63. Zikmund-Fisher BJ. The right tool is what they need, not what we have: A taxonomy of appropriate levels of precision in patient risk communication. *Med Care Res Rev*. 2013;70:37S–49S.
64. Rao G. Physician numeracy: Essential skills for practicing evidence-based medicine. *Fam Med*. 2008;40:354.
65. Schulkin J, Anderson BL. Physicians' understanding and use of numeric information. In: Anderson BL, Schulkin J eds. *Numerical Reasoning in Judgments and Decision Making about Health*. Cambridge: Cambridge University Press; 2014.
66. Caverly TJ, Hayward RA, Reamer E, et al. Presentation of benefits and harms in US cancer screening and prevention guidelines: Systematic review. *J Natl Cancer Inst*. 2016;108: djv436.
67. Schwartz LM, Woloshin S, Welch HG. The drug facts box: Providing consumers with simple tabular data on drug benefit and harm. *Med Decis Making*. 2007;27:655–62.
68. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. Improving understanding of adjuvant therapy options by using simpler risk graphics. *Cancer* 2008;113:3382–90.
69. Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. *Cochrane Database Syst. Rev*. 2013;1: CD004720.
70. Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst. Rev*. 2013;6: CD001877.