The Effectiveness of a Primer to Help People Understand Risk
Two Randomized Trials in Distinct Populations

Steven Woloshin, MD, MS; Lisa M. Schwartz, MD, MS; and H. Gilbert Welch, MD, MPH

Background: People need basic data interpretation skills to understand health risks and to weigh the harms and benefits of actions meant to reduce those risks. Although many studies document problems with understanding risk information, few assess ways to teach interpretation skills.

Objective: To see whether a general education primer improves patients' medical data interpretation skills.

Design: Two randomized, controlled trials done in populations with high and low socioeconomic status (SES).

Setting: The high SES trial included persons who attended a public lecture series at Dartmouth Medical School, Hanover, New Hampshire; and the low SES trial included veterans and their families from the waiting areas at the White River Junction Veterans Affairs Medical Center, White River Junction, Vermont.

Participants: 334 adults in the high SES trial and 221 veterans and their families in the low SES trial were enrolled from October 2004 to August 2005. Completion rates for the primer and control groups in each trial were 95% versus 98% (high SES) and 85% versus 96% (low SES).

Intervention: The intervention in the primer groups was an educational booklet specifically developed to teach people the skills needed to understand risk. The control groups received a general health booklet developed by the U.S. Department of Health and Human Services Agency for Health Care Research and Quality.

Measurements: Score on a medical data interpretation test, a previously validated 100-point scale, in which 75 points or more is considered "passing." Secondary outcomes included 2 other 100-point validated scores (interest and confidence in interpreting medical statistics) and participants' ratings of the booklet's usefulness.

Results: In the high SES trial, 74% of participants in the primer group received a "passing grade" on the medical data interpretation test versus 56% in the control group ($P = 0.001$). Mean scores were 81 and 75, respectively ($P = 0.0006$). In the low SES trial, 44% versus 26% "passed" ($P = 0.010$): Mean scores were 69 and 62 in the primer and control groups, respectively ($P = 0.008$). The primer also significantly increased interest in medical statistics by 6 points in the high SES trial (a 4-point increase vs. a 2-point decrease from baseline) ($P = 0.004$) and by 8 points in the low SES trial (a 6-point increase vs. a 2-point decrease from baseline) ($P = 0.004$) compared with the control booklet. The primer, however, did not improve participants' confidence in interpreting medical statistics beyond the control booklet (a 2-point vs. a 4-point increase in the high SES trial [$P = 0.36$] and a 2-point versus a 6-point increase in the low SES trial [$P = 0.166$]). The primer was rated highly: 91% of participants in the high SES trial found it "helpful" or "very helpful," as did 95% of participants in the low SES trial.

Limitations: The primarily male low SES sample and the primarily female high SES sample limits generalizability. The authors did not assess whether better data interpretation skills improved decision-making.

Conclusion: The primer improved medical data interpretation skills in people with high and low SES.

People face a bewildering array of medical decisions (for example, should I be tested for BRCA1? Should I be screened for prostate cancer? Should I have lumpectomy and radiation to treat my breast cancer?). If people are to make informed decisions, they need to understand risk: What is my chance of staying healthy if I undergo this intervention? What is my chance of staying healthy if I forgo it? What is the chance of harm?

Understanding risk data entails a set of skills. First, people must be able to work with probabilities and changes in probabilities, the typical measures used to communicate risk information. Next, they need to have a framework on which to organize data. In essence, this means having a sense of what additional information is needed to give meaning to a statement about probability (conversely, how to know when such additional information is lacking). For example, putting a particular risk in context ("your chance of breast cancer is 1 in 8") entails knowing what outcome is being considered (diagnosis vs. death), being clear about the time period (5 years vs. lifetime), and having some perspective on the magnitude of the probability (How does

the risk for a particular condition compare with other important health risks?). Third, people need to be able to make some basic assessment about the quality of the evidence to know whether they can believe the numbers that they are given.

Anecdotal experience and a growing body of literature document the trouble many people have in making sense of risk information (1–4), which is now ubiquitous in health messages. This is not surprising: Much more effort goes into disseminating health information (for example, drug advertisements, media reports, and decision-making aids) than into preparing the target audiences to under-

See also:

Print

Web-Only
Conversion of figures and tables into slides

stand it. To address this gap, we developed a primer to help people develop the basic skills needed to make sense of the medical risk data that surround them. In this paper, we report on 2 randomized trials that tested the primer. We sought to learn how people would respond to the primer: Would they read it? Would they value this knowledge? Most important, would the primer improve patients' abilities to interpret medical data?

METHODS

Design Overview

To test the effect of the primer on how well people understand risk, we conducted 2 randomized trials in distinct populations: people with high and low socioeconomic status (SES). **Figure 1** shows an overview of the study design. The Committee for the Protection of Human Subjects at Dartmouth College approved this project, and the survey cover letter was considered to be informed consent (signed consent was waived).

Setting and Participants

We calculated our sample size under the most conservative conditions—when the pooled proportion of passing scores was 50%. We asserted that a 20-percentage point absolute difference in the proportion passing the test in the primer group versus the control group would be clinically important. In calculating the sample size requirements, we therefore assumed that the proportion passing would be 50% in the control group and 70% in the primer group. When a power of 0.8 and a 2-sided P value of 0.05 were used, 100 patients were required for each study group. Assuming that 10% of participants would be lost to follow-up, we planned to enroll 110 patients per group in each trial.

High Socioeconomic Status Trial

To study the effect of the primer in a highly educated and affluent group, we recruited alumni from Dartmouth's "Community Medical School." This 9-lecture series on various health and medical topics is taught by Dartmouth faculty and guests; is held annually in Hanover, New Hampshire, and Manchester, New Hampshire; and has a $25 registration fee (5). For this study, the program organizers mailed recruitment letters to 1138 "alumni." A total of 334 people who responded to the letter were eligible (that is, they met the age criterion of 35 to 79 years, spoke English, and had attended the Community Medical School in the past) and were subsequently randomly assigned. Ninety-six percent ($n = 322$) returned a completed survey (completion rates were not significantly different between the primer and control groups [95% vs. 98%, respectively]) ($P = 0.192$).

Low Socioeconomic Status Trial

To study the effect of the primer in a sample with lower income and less formal education, we recruited veterans and their families at the White River Junction Vet-

Context

Educational materials aimed at improving people's ability to understand information about risk are scarce.

Contribution

In 2 trials, adults with high or low socioeconomic status (SES) were randomly assigned to receive a primer about understanding risk or a general health booklet. In both SES groups, adults receiving the primer more often passed a medical data interpretation test than did those receiving the general health booklet. They also expressed greater interest in medical statistics but not greater confidence in interpreting statistics, and most rated the primer helpful or very helpful.

Cautions

The authors did not examine whether improved data interpretation skills affected decision-making.
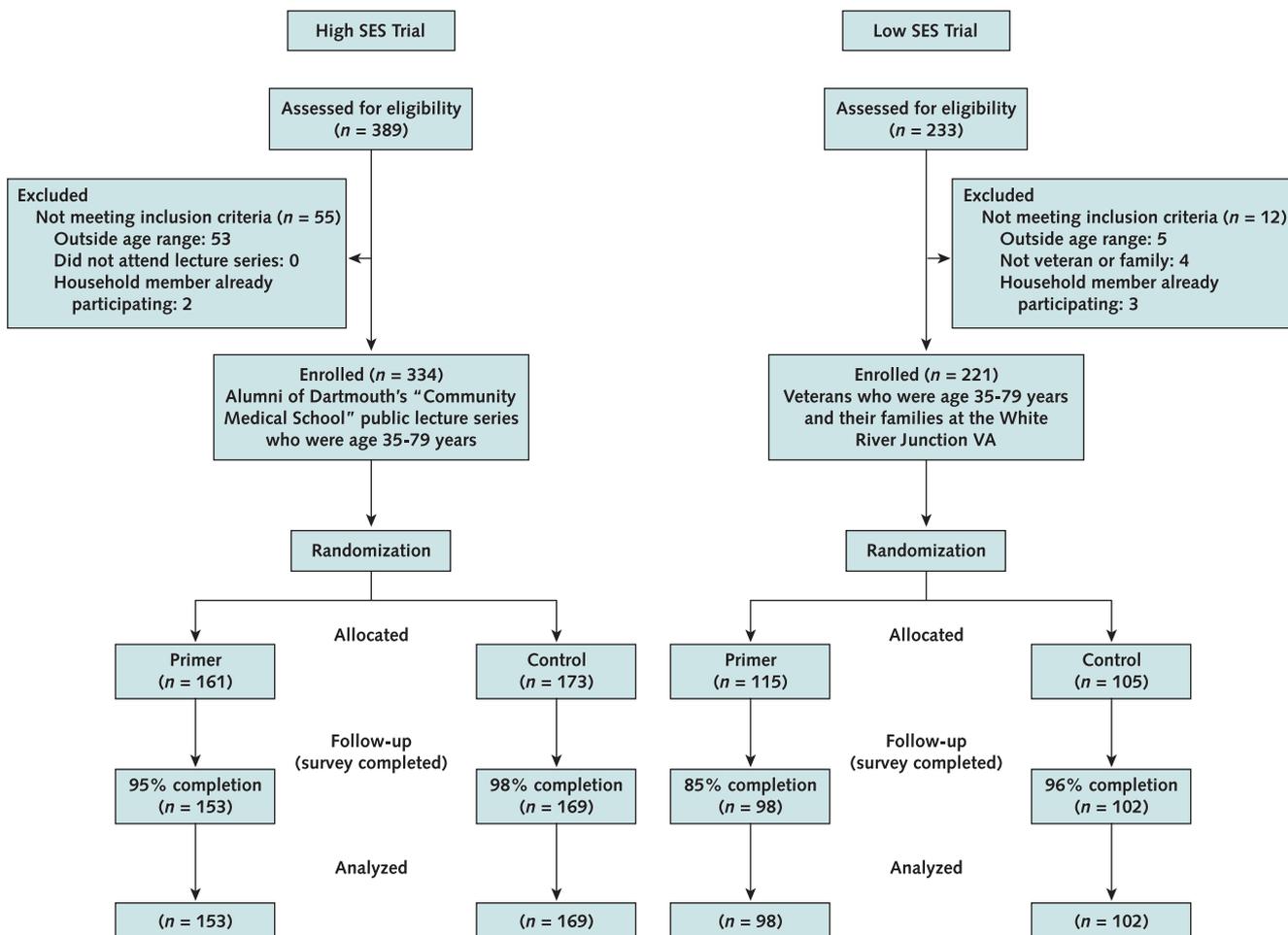
—The Editors

erans Affairs (VA) Medical Center, White River Junction, Vermont, by posting study advertisements in waiting areas of the outpatient clinic. A total of 221 people who responded to the advertisement were eligible (that is, they met the age criteria of 35 to 79 years, spoke English, and were a veteran or the family member of a veteran who was enrolled in a VA clinic) and were subsequently randomly assigned. Ninety percent ($n = 200$) returned completed surveys (completion rates were significantly lower in the primer group than in the control group [85% vs. 96%, respectively]) ($P = 0.005$).

Randomization and Interventions

The letter and the advertisement asked people to participate in a research study to learn how to better give people health information. We did not mention our interest in enhancing quantitative skills. People who responded to the recruitment letter (high SES group) or to the advertisement (low SES group) were first interviewed to confirm eligibility. We only accepted 1 participant per household. Within each trial, we assigned participants on an individual basis to receive either the primer or control booklet. A list of random numbers (created by using a random-number generator) was given to a research assistant and was used to determine assignments (randomization was not stratified or blocked within each trial). The research assistant had access to participants' characteristics. The investigators did not have access to the assignment list, and the deidentified code was only revealed after recruitment and completion of data collection.

The appropriate booklet and a survey were mailed or given in person to participants. The survey included the major outcomes that will be described in the next section. To make the survey seem relevant to the control group, it

Figure 1. Overview of 2 randomized trials in distinct populations.

High SES Trial

Assessed for eligibility
(*n* = 389)

Excluded
Not meeting inclusion criteria (*n* = 55)
Outside age range: 53
Did not attend lecture series: 0
Household member already
participating: 2

Enrolled (*n* = 334)
Alumni of Dartmouth's "Community
Medical School" public lecture series
who were age 35–79 years

Randomization

Allocated

Primer
(*n* = 161)

Control
(*n* = 173)

Follow-up
(survey completed)

95% completion
(*n* = 153)

98% completion
(*n* = 169)

Analyzed

(*n* = 153)

(*n* = 169)

Low SES Trial

Assessed for eligibility
(*n* = 233)

Excluded
Not meeting inclusion criteria (*n* = 12)
Outside age range: 5
Not veteran or family: 4
Household member already
participating: 3

Enrolled (*n* = 221)
Veterans who were age 35–79 years
and their families at the White
River Junction VA

Randomization

Allocated

Primer
(*n* = 115)

Control
(*n* = 105)

Follow-up
(survey completed)

85% completion
(*n* = 98)

96% completion
(*n* = 102)

Analyzed

(*n* = 98)

(*n* = 102)

SES = socioeconomic status; VA = Veterans Affairs.

also included additional questions about facts presented in the health booklet (we did not analyze responses to these questions). All participants were asked to read the booklet and return the completed survey within 2 weeks using the self-addressed stamped envelope included with the survey. Reminder letters were sent to nonresponders. Participants who returned surveys were given their choice of a $25 gift certificate to a local bakery, restaurant, bookstore, or large retail store. We recruited participants from October 2004 through August 2005.

Primer Group

The participants in the primer group received the booklet entitled "Know Your Chances: Understanding Health Statistics." The goal of the primer is to teach people how to understand risk messages and health statistics. The first part of the primer teaches people how to understand disease risk by using the example of colon cancer. The second part focuses on how to understand the benefits and

harms of interventions by using a Zocor direct-to-consumer advertisement for secondary heart disease prevention. **Figure 2** shows the summary of main concepts that were taught, which is excerpted from the booklet. The contents of the primer and many of the examples have been developed and revised over years of teaching and through focus groups with people across a diverse socioeconomic spectrum.

Because many people are intimidated by numbers and statistics, we worked hard to make the primer inviting and nonthreatening by liberal use of cartoons and figures; by working through examples, separating the most technical material into optional "learn more boxes"; and by providing readers with quizzes (with answers) to assess their mastery of the material as they read the primer. **Figure 3** shows sample pages from the primer. Most of the primer is written at the eighth-grade or lower reading level (6) and is a color document approximately 80 pages in length.

Table 1. **Baseline Characteristics of Participants in the High Socioeconomic Status and Low Socioeconomic Status Trials***

Variable	High SES Trial (*n* = 322)		Low SES Trial (*n* = 200)	
	Control Group (*n* = 169)	Primer Group (*n* = 153)	Control Group (*n* = 102)	Primer Group (*n* = 98)
Mean age, *y* (range)	60 (41–78)	61 (41–76)	58 (37–79)	57 (38–76)
Women, %	85	76	20	19
Household income, %				
<$10 000	1	1	9	13
$10 000–$24 999	4	4	31	23
$25 000–$49 999	18	24	34	33
$50 000–$99 999	48	48	23	26
≥$100 000	30	22	3	5
Highest level of education, %				
<High school graduate	0	0	6	3
High school graduate	18	21	47	50
College degree	42	42	33	30
Postgraduate degree	40	37	13	17
Race/ethnicity, %				
White	99	100	91	93
Black	0	0	2	3
Hispanic	1	0	3	3
Other	0	0	4	1
Health status, %				
Excellent	22	32	6	6
Very good	51	39	33	22
Good	21	23	33	43
Fair	7	5	23	17
Poor	0	1	5	11
Medical conditions, %				
Heart or vascular disease	8	6	19	24
COPD	1	0	9	6
Diabetes	5	6	24	19
History of any type of cancer	21	19	20	23
Current smoker, %	3	0	15	16
Mean score on attitudes toward medical statistics†				
Interest	58	57	59	59
Confidence interpreting	65	65	65	62

* The results are for completers only. Percentages may not add to 100% because of rounding. Item nonresponse for high and low SES trials, respectively, were age, 1% and 1%; sex, 0% and 0%; income, 6% and 4%; education, 0.4% and 0.5%; race, 0.4% and 0.5%; health status, 1% and 1%; medical conditions, 0% and 0%; smoking, 0% and 0.5%; interest in statistics, 2% and 3%; and confidence, 0.7% and 0.5%. COPD = chronic obstructive pulmonary disease; SES = socioeconomic status.
† Scored on scales of 0 to 100 where higher scores represent higher interest or confidence.

Control Group

The control group received a 70-page booklet entitled "The Pocket Guide for Good Health for Adults," which is published by the U.S. Department of Health and Human Services Agency for Health Care Research and Quality (7). We chose this booklet because the length is similar to that of the primer, it is written at a similar reading level (eighth grade or lower) (6), and it contains general information about risk and reducing risk by following recommended prevention or screening activities. However, it does not include training on how to interpret quantitative information.

Measurement and Outcomes

All outcome measures were assessed in the survey.

Primary Outcome

The primary outcome was a measure of participants' abilities to interpret medical statistics. To measure this outcome, we developed the 18-item data interpretation test during the same time as the primer and validated it in a separate study (8). The test is available at www.vaoutcomes.org/research_tools.php. It extends our original 3-item numerical measure (4) to include a much broader set of skills beyond simple mathematical manipulation: such skills are the ability to compare risks, to put risk estimates into con-

Figure 2. Key concepts taught in the primer, excerpted from the final pages.

Questions to ask when interpreting risk	Questions to ask when interpreting benefit
Risk of what? Understand what the outcome is (getting disease, dying from disease, developing a symptom) and consider how bad it is. *How big is the risk?* Understand the chance that you will experience the outcome. When you hear about the number of people with the disease (or whatever outcome you are discussing), always ask, "Out of how many?" so you can learn the chance of the disease. Also determine the time period of the risk (over the next year, 10 years, a lifetime). *Does the risk information reasonably apply to you?* Understand whether the message is based on people like you (people of your age and sex, people whose health is like yours). *How does this risk compare to other risks?* Ask for some context so you can develop a sense of just how big (or small) the risk really is.	*Change in risk of what?* Understand what the outcome is (getting disease, dying from disease, developing a symptom) and consider how bad it is. *How big is the change in risk?* Whenever you hear about changing risk (e.g., "42% lower"), ask "Lower than what?" Learn what your risk is if you do something (e.g., take a drug, change your lifestyle) versus if you don't. *Does the change in risk information reasonably apply to you?* Understand whether the people the message is based on are like you (people of your age and sex, people whose health is like yours). *How does this change in risk compare to other changes in risk?* Ask for some context so you can develop a sense of just how big (or small) the change in risk really is.
Things you should do to better understand risk *Get the risk information into a user-friendly format.* Make the numbers as easy as you can by translating "1 in __" to "__ out of 1000." *Try framing the risk in different ways.* If they tell you the chance of dying, rewrite it as the chance of living to see if this changes how the information makes you feel.	**Things you should do to better understand benefit** *Get the bottom line.* Learn what your overall chance of dying is if you do something versus if you don't. *Learn the downsides (e.g. possible side effects) that come with the benefit.* Understand what the side effects are, how often they happen, and how bad they are. *Find out the source of information, the kind of research, and what is being measured.* Be wary of information from scientific meetings, from sources with other interests besides your health (like money), from animal research or studies that just describe what happens to people.

text, and to know what additional information is necessary to give meaning to a medical statistic. The data interpretation test does not test recall of facts; instead, it asks people to interpret real-world information (for example, drug advertisements and news stories). The skills being tested are the same as those in the primer because we feel that they are key skills for people to understand. However, the medical data interpretation test requires that the respondent apply these skills to data (within various messages) that do not appear in the primer.

Scores from this test range from 0 to 100—higher scores represent higher abilities. We consider a score of 75 or higher as "passing" and a score of 90 or higher as "outstanding" (the latter threshold corresponds to the mean score of physician experts who teach "evidence-based" medicine in our validation study [8]).

Secondary Outcomes

Secondary outcomes included measures of interest and confidence in interpreting medical statistics. These outcomes were assessed twice: before randomization (at the eligibility interview) and again as part of the survey. These previously validated measures (STAT-interest and STAT-confidence scales) are also scored on 0 to 100 scales in which higher scores represent higher interest or confidence (9). In addition, we asked participants to rate each booklet on various dimensions (for example, Did you find it helpful? Did you learn something new?, and Would you recommend it to others?).

Statistical Analysis

The main results were based on complete case analysis with additional sensitivity analyses to address the impact of missing data. We used the chi-square test to compare differences in proportions and *t*-tests to compare means. All comparisons were 2-sided and were considered statistically significant at *P* values less than 0.05. We used Stata, version 9.1 (Stata Corp., College Station, Texas) for all analyses.

Role of the Funding Sources

Drs. Woloshin and Schwartz are supported by Robert Wood Johnson Generalist Faculty Scholars Awards and receive infrastructure support from a Research Enhance-

ment Award from the Department of Veterans Affairs. The study was supported by a grant from the National Cancer Institute (R01CA104721). The funding sources did not play a role in the design, data collection, analysis, or interpretation of the study. The authors had full access to the data files for this study.

RESULTS

Participants

Table 1 summarizes the characteristics of the participants who completed the survey. As designed, the 2 trials had different participants. In the high SES trial, 70% had

a household income of $50 000 or greater and 80% had a college or postgraduate degree. In the low SES trial, approximately 70% had a household income less than $50 000 and approximately 50% had a high school degree or less formal education. As expected, participants in the low SES trial were mostly men (recruitment was done at a VA hospital) and the participants in the high SES trial were mostly women (as has typically been the case at the Dartmouth Community Medical School). The participants in the low SES trial had lower self-reported health status, were more likely to smoke, and had more medical conditions than those in the high SES trial. Within each trial, however, there were no statistically signifi-

Table 2. Participants' Ratings of the Primer and Control Booklets*

Booklet Use	High SES Trial			Low SES Trial		
	Control Group (n = 169), %	Primer Group (n = 153), %	P Value	Control Group (n = 102), %	Primer Group (n = 98), %	P Value
About how much time did you spend reading the booklet?			<0.001			0.001
≤10 min	7	1		4	2	
30 min	65	27		48	24	
≥1 h	28	72		47	74	
Was the booklet easy or hard to understand?			<0.001			0.003
Very easy	81	42		60	37	
Easy	19	55		39	53	
Hard	0	3		1	9	
Very hard	0	0		0	1	
Value of booklet						
How helpful was the booklet?			<0.001			0.257
Very helpful	16	51		41	37	
Helpful	53	40		49	57	
A little	30	9		11	5	
Not at all	1	0		0	0	
I plan to refer to the booklet when I make medical decisions in the future			<0.001			0.217
Definitely yes	10	28		25	27	
Probably yes	41	52		57	47	
Probably no	44	14		17	21	
Definitely no	5	5		1	5	
I will recommend the booklet to others			<0.001			0.233
Definitely yes	26	45		50	38	
Probably yes	43	42		40	52	
Probably no	26	11		9	6	
Definitely no	5	3		2	4	
I wish I had read the booklet before			<0.001			0.302
Definitely yes	8	22		24	34	
Probably yes	34	49		52	41	
Probably no	49	23		22	22	
Definitely no	9	6		2	3	
I learned something new			<0.001			0.096
Definitely yes	28	54		51	68	
Probably yes	38	37		39	24	
Probably no	26	8		8	5	
Definitely no	7	1		2	3	

* The results are for completers only. Completion rates for the control and primer groups were as follows: 98% vs. 95% (high SES trial) and 96% vs. 85% (low SES trial). Percentages may not add to 100% because of rounding. Item nonresponse for the high and low SES trials, respectively, were time, 2.3% and 2.5%; difficulty, 0.4% and 3%; helpfulness, 1% and 2.5%; refer to later, 1.9% and 1.5%; recommend to others, 1.6% and 2%; read before, 1.9% and 1.5%; and learned, 1% and 1.5%. SES = socioeconomic status.

Figure 3. Excerpts of primer.

Sample Pages from Part I: What Is My Risk?

Chapter 1
Interpreting Risk

The early warning signs of colon cancer:

You feel great.

You have a healthy appetite.

You're only 50.

[Sloan Kettering ad, NYT]

For many people this is a pretty scary message. It says, worry: If you feel well, you may have colon cancer (cancer of the large intestine). To be fair, the purpose of the message is to get people to go for colon cancer screening. But some key information is missing from this message. For starters, it doesn't tell you how likely colon cancer is. If your chance is big, there is more reason to worry than if your chance is small. This information might help you decide whether to try to lower your risk.

QUIZ

Do all 50-year-old people who feel great and have healthy appetites really have colon cancer?

a. Yes

b. No

3

KNOW YOUR CHANCES

The correct answer is the total number of women. (That is, 12,000 out of *how many?*) Here is a more complete way to talk about colon cancer risk:

"Colon cancer will strike 1 in 21 people."

LEARN MORE

To calculate a risk, you need to know two things: how many people have experienced something (the numerator) and how many people could experience it (i.e., the denominator).

$$Risk = \frac{People\ with\ colon\ cancer\ (numerator)}{People\ at\ risk\ (denominator)} = \frac{135,000}{275,000,000\ Americans} = 0.05\%$$

6

Sample Pages from Part II: What Is the Benefit?

KNOW YOUR CHANCES

According to the ad, the man in the picture had a heart attack, then found out he had high cholesterol. He is worried about the future and wants to reduce his risk. The ad says that the medicine ZOCOR can help him. The white oval highlights the main message about how ZOCOR will change his heart risk. What does it mean?

"A clinical study among people with high cholesterol and heart disease found 42% fewer deaths from heart attack among those taking ZOCOR."
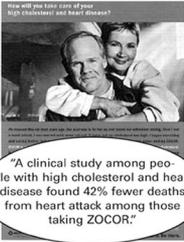
As you know from the last chapter, the first questions to ask about a risk are: Risk of what? and How big is the risk? The same questions apply to changing risk. The following quiz will help to get us started answering these questions.

QUIZ

Does ZOCOR change the chance of having or dying from a heart attack?

a. Having a heart attack

b. Dying from a heart attack

How much does ZOCOR change the risk?

a. From 1000 in 1000 to 580 in 1000

b. From 10 in 1000 to 5.8 in 1000

c. Can't tell

34

KNOW YOUR CHANCES

Thinking about changing risk is like deciding when to use a coupon at a store. Imagine that you have a coupon that says "50% off any one purchase." You are going to the store to buy a pack of gum, which costs 50 cents, and a large Thanksgiving turkey for $35. Will you use the coupon for the gum or the turkey? Most people would use the coupon for the turkey. Here's why:

| ITEM	REGULAR PRICE	SALES PRICE	SAVINGS
Gum	$0.50	$0.25	$0.25
Turkey	$35.00	$17.50	$17.50

In both cases, you saved 50%. But with a cheap item, that means little (25 cents); with an expensive item, that means a lot ($17.50). The bottom line: To know what a discount means, you need to know the regular price. The same thing in health care.

That turkey! He used his coupon on the gum!

uh oh! my beak is stuck!

QUIZ

Imagine that you are a typical 70-year-old woman who has never smoked. Consider 2 drugs. Both drugs lower the chance of dying from a disease by 50%. However, one works on cervical cancer and the other works on heart attacks. Which drug is more likely to help you? HINT: Refer to the risk charts in the Extra Help section.

a. Drug for cervical cancer

b. Drug for heart attack

37

Figure 4. Medical data interpretation scores (mean and proportion exceeding threshold) for the primer and control groups.
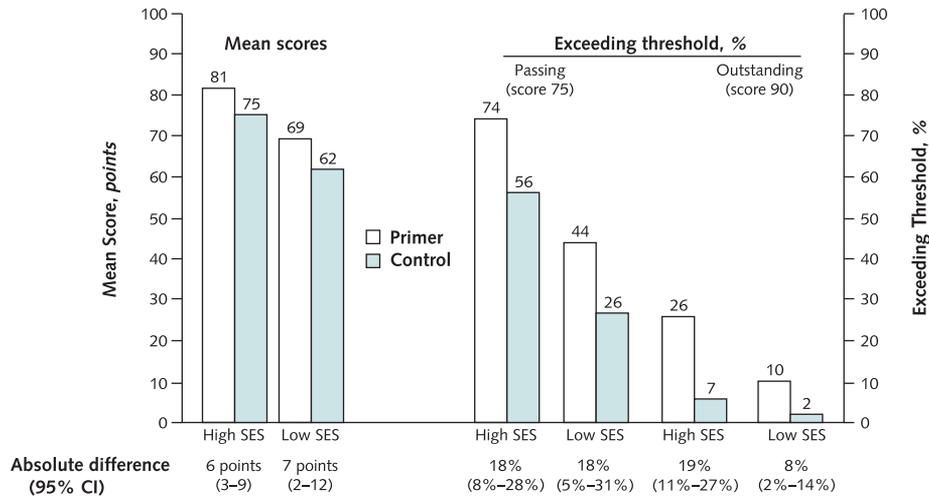
SES = socioeconomic status.

Primary Outcome: Medical Data Interpretation Skills

In both trials, the primer resulted in higher medical data interpretation scores than did the control booklet (**Figure 4**). In the high SES trial, the mean score was 81 in the primer group versus 75 in the control group—a 6-point difference ($P = 0.0006$). In terms of score thresholds, 74% of participants in the primer group received a "passing grade" (score ≥75) versus 56% of participants in the control group ($P = 0.001$). Twenty-six percent of participants in the primer group received an "outstanding grade" (score of ≥90), which is the mean score of physician experts who teach evidence-based medicine in a validation study of this scale versus 7% of participants in the control group ($P < 0.001$). In the low SES trial, the mean scores were 69 in the primer group versus 62 in the control group—a 7-point difference ($P = 0.008$). Passing grades were 44% versus 26% ($P = 0.010$) and outstanding grades were 10% versus 2% in the primer and controls groups, respectively ($P = 0.014$). Finally, results were similar in additional analyses that adjusted for age, sex, level of education, and medical conditions: The difference in mean interpretation scores was 7 points ($P < 0.001$) in the high SES trial and 6 points ($P = 0.019$) in the low SES trial for the primer versus the control groups, respectively.

The specific skills for which participants in the primer group most outperformed those in the control group were as follows: recognizing that counts of events without denominators do not convey risks (77% vs. 59% in the high SES trial and 54% vs. 41% in the low SES trial answered correctly) and knowing that a risk statement applies to you requires knowing about the age and sex in the source data for the risk (81% vs. 55% in the high SES trial and 62% vs. 37% in the low SES trial answered correctly).

To assess the effect of nonresponse on our findings, we considered various possibilities of pass rates among the nonresponders. Assuming that nonresponders have lower data interpretation skills, the most plausible "worst-case" scenario is that the nonresponders would have failed the test. In this case, the statistical significance of the low SES trial becomes borderline (37% vs. 25%; $P = 0.057$) but remains significant in the high SES trial (70% vs. 54%). In the least plausible and most extreme "worst-case" scenario (assuming that all of the nonresponders in the primer group would have failed the test but that all nonresponders in the control group would have passed the test), the difference in pass rates in the low SES trial becomes nonsignificant (37% vs. 29%; $P = 0.20$), but remains significant in the high SES trial (70% vs. 57%).

Secondary Outcomes

Interest and Confidence

Interest in medical statistics was significantly higher (6 points) in the high SES trial (a 4-point increase vs. 2-point decrease from baseline; $P = 0.004$) and by 8 points in the low SES trial (a 6-point increase vs. a 2-point decrease from baseline; $P = 0.004$) after exposure to the primer compared with the control booklet. The primer, however, did not improve participants' confidence in interpreting medical statistics beyond the control booklet; if anything, confidence improved more with the control booklet: by 2 points in the high SES trial (a 2-point increase vs. a 4-point increase; $P = 0.36$) and by 4 points in the low SES trial (a 2-point increase vs. a 6-point increase; $P = 0.166$).

Rating the Booklets

Finally, participants spent substantial time reading the primer and rated it highly (**Table 2**). In the high SES trial, 72% of the primer group reported spending an hour or more reading the primer, whereas only 28% of the control

group spent that much time reading their booklet ($P <$ 0.001). Similarly, in the low SES trial, 74% of the primer group spent an hour or more reading the primer, whereas 47% of the control group spent this much time reading their booklet ($P = 0.001$).

In the high SES trial, all ratings were significantly higher for the primer group than for the control group. For example, 91% found the primer "helpful" or "very helpful" versus 69% for the control booklet ($P < 0.001$). In the low SES trial, the primer ratings were high but were not statistically different from the ratings of the control booklet. For example, 94% and 90% of participants rated the primer and control booklet "helpful" or "very helpful," respectively ($P = 0.26$).

No analyses, including those for the secondary outcomes and booklet ratings, include participants who did not complete the survey. Because completion rates were lower in the primer group versus the control group in the low SES trial (85% vs. 96%), the secondary outcomes and booklet ratings may favor the primer group to the extent that nonresponders had a less favorable response to the primer.

DISCUSSION

We developed and tested a general education primer designed to teach people basic medical data interpretation skills. The primer improved these skills in 2 distinct populations: participants with high SES and a demonstrated interest in health information (that is, they had attended a "community medical school" lecture series) and those with generally low formal education and limited quantitative skills. Participants liked the primer: A substantial majority of participants in both trials said they found it useful, it taught them something new, and that they would recommend it to others.

Our findings have several limitations. First is the question of generalizability. Our study samples were chosen deliberately to ensure the participation of people across a broad spectrum of age, income, and formal education. Because almost all participants were white and English speakers, we cannot be certain how the primer would perform in a more ethnically diverse setting. Also, because all participants were paid—those recruited from an actual health care setting and those recruited from the community—we cannot be certain how unpaid persons would respond to the primer. Second, there may be concern regarding the appropriateness of the control booklet. We chose an educational booklet (7) (published by the U.S. Department of Health and Human Services Agency for Healthcare Research and Quality) that was similar to the primer in length and tone. Because the control booklet did not teach readers about interpreting risk, it worked as a "placebo" with respect to the main outcome measure (that is, the tests results on medical data interpretation in the control group should reflect existing abilities). Third, in both trials, completion rates were lower in the primer group than in the control group, particularly in the low SES trial. This observation is

not hard to explain: The primer covered more challenging material (it required readers to do math) than did the control book. However, completion rates were high overall (95% vs. 98% in the high SES trial and 85% vs. 97% in the low SES trial).

Curiously, although the primer improved ability and stimulated general interest in medical statistics, it did not increase participants' confidence in interpreting these statistics any more than did the control booklet. This finding probably reflects the educational impact of the primer. For some people (especially those with the least formal education), acquisition of new data interpretation skills may highlight what they did not know before and may expose overconfidence.

How big was the change in data interpretation skills? Because readers may be unfamiliar with our main outcome measure, they may have trouble gauging the clinical significance of our results. One way to put our findings into perspective is to consider the distribution of medical data interpretation skills observed in our previous validation study. In that study, the average difference in scores between experts (medical school faculty who teach evidence-based medicine) and other respondents with postgraduate degrees (similar educational attainment but no relevant expertise) was 20 points (8). The 6- and 7-point differences observed in the present study correspond to gaining approximately one third of the difference. Another way to put our findings into perspective is to consider the magnitude of the primer's effect in relation to other educational interventions. One way to make comparisons across interventions is to translate our findings into generic "effect size" units by using a metric called Cohen's d (primer minus control medical data interpretation scores divided by the pooled standard deviation of the scores) (10). Our results correspond to Cohen's d effect sizes of 0.38 (high SES) and 0.42 (low SES); by convention, these effect sizes are considered to be between "small" and "medium," the range covering many successful educational interventions (11). For context, on average, taking a practice examination is associated with an approximate 0.30-effect size unit increase on standardized test scores (12).

Finally, it is important to acknowledge that although we have demonstrated that exposure to the primer improved medical data interpretation skills, we did not prove that the primer leads to better decision-making. Because informed decision-making is not possible without understanding the relevant facts, we think improving data interpretation skills is clearly a legitimate outcome in itself and a necessary prerequisite to making good decisions.

Implementing the primer, like decision aids, guidelines, or research innovations in general (13) may be challenging. We believe there are several ways that the primer could be used in clinical practice. It might be distributed either before or after routine clinic visits when patients are making decisions regarding screening and prevention. It might be even more important for patients facing such

high-stakes decisions as surgery or chemotherapy. It is possible that patients facing a new and serious disease (for example, a new diagnosis of cancer) may have greater motivation to really understand the data regarding benefits and harms of various interventions and would find the primer very useful. Alternatively, they may feel too emotionally overwhelmed to use the primer. If this were the case, the primer might be useful for families or friends who are helping the patient with decision-making. Questions of when and where the primer might be most effectively used and how the reader's emotional state, decision-making style, and preferences might influence its impact are clearly a topic for further study.

In conclusion, the primer "Know Your Chances: Understanding Health Statistics"—a simple, inexpensive, low-tech intervention—improved medical data interpretation skills and was rated highly by well-educated participants as well as by those with less formal education.

References

1. **Edwards A, Elwyn G.** Understanding risk and lessons for clinical risk communication about treatment preferences. Qual Health Care. 2001;10 Suppl 1:i9-13. [PMID: 11533431]

2. **Gazmararian JA, Baker DW, Williams MV, Parker RM, Scott TL, Green DC, et al.** Health literacy among Medicare enrollees in a managed care organization. JAMA. 1999;281:545-51. [PMID: 10022111]

3. **Kirsch IS, Jungeblut A, Jenkins L, Kolstad A.** Adult literacy in America: A First Look at Results of the National Adult Literacy Survey. Washington, DC: U.S. Government Printing Office; 1993.

4. **Schwartz LM, Woloshin S, Black WC, Welch HG.** The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med. 1997;127:966-72. [PMID: 9412301]

5. Dartmouth Community Medical School Home Page. Accessed at http://dms.dartmouth.edu/dcms on 14 December 2006.

6. **Microsoft Corporation.** Microsoft Word X. Flesch-Kincaid Readability Statistics. Redmond, WA: Microsoft Corporation; 1983-2001.

7. **Agency for Healthcare Research and Quality.** The pocket guide for good health for adults. Accessed at www.ahrq.gov/ppip/adguide/ on 14 September 2006.

8. **Schwartz LM, Woloshin S, Welch HG.** Can patients interpret health information? An assessment of the medical data interpretation test. Med Decis Making. 2005;25:290-300. [PMID: 15951456]

9. **Woloshin S, Schwartz LM, Welch HG.** Patients and medical statistics. Interest, confidence, and ability. J Gen Intern Med. 2005;20:996-1000. [PMID: 16307623]

10. **Thalheimer W, Cook S.** How to calculate effect sizes from published research articles: a simplified methodology. Accessed at www.work-learning.com/effect_sizes.htm on 14 September 2006.

11. **Valentine J, Cooper H.** Effect size substantive interpretation guidelines: Issues in the interpretation of effect sizes. Washington, DC: What Works Clearinghouse; 2003. Accessed at www.whatworks.ed.gov/reviewprocess/essig.pdf on 15 September 2006.

12. **Kulik JA, Kulik CLC, Bangert RL.** Effects of practice on aptitude and achievement test scores. Am Educ Res J. 1984;21:435-47. Accessed at www.jstor.org/view/00028312/ap040083/04a00130/0 on 14 December 2006.

13. **Haines A, Jones R.** Implementing findings of research. BMJ. 1994;308:1488-92. [PMID: 8019284]

Annals of Internal Medicine

Current Author Addresses: Drs. Woloshin, Schwartz, and Welch: VA Outcomes Group (11B), Department of Veterans Affairs Medical Center, White River Junction, VT 05009.

Author Contributions: Conception and design: S. Woloshin, L.M. Schwartz, H.G. Welch.
Analysis and interpretation of the data: S. Woloshin, L.M. Schwartz, H.G. Welch.

Drafting of the article: S. Woloshin, L.M. Schwartz, H.G. Welch.
Critical revision of the article for important intellectual content: S. Woloshin, L.M. Schwartz.
Final approval of the article: S. Woloshin, L.M. Schwartz, H.G. Welch.
Provision of study materials or patients: S. Woloshin, L.M. Schwartz.
Statistical expertise: S. Woloshin, L.M. Schwartz.
Obtaining of funding: S. Woloshin, L.M. Schwartz.