Patients and Medical Statistics

Interest, Confidence, and Ability

Steven Woloshin, MD, MS,[1,2] Lisa M. Schwartz, MD, MS,[1,2] H. Gilbert Welch, MD, MPH[1,2]

[1]The VA Outcomes Group (111B), Department of Veterans Affairs Medical Center, White River Junction, Vt, USA; [2]The Center for the Evaluative Clinical Sciences, Dartmouth Medical School, Hanover, NH, USA.

BACKGROUND: People are increasingly presented with medical statistics. There are no existing measures to assess their level of interest or confidence in using medical statistics.

OBJECTIVE: To develop 2 new measures, the STAT-interest and STAT-confidence scales, and assess their reliability and validity.

DESIGN: Survey with retest after approximately 2 weeks.

SUBJECTS: Two hundred and twenty-four people were recruited from advertisements in local newspapers, an outpatient clinic waiting area, and a hospital open house.

MEASURES: We developed and revised 5 items on interest in medical statistics and 3 on confidence understanding statistics.

RESULTS: Study participants were mostly college graduates (52%); 25% had a high school education or less. The mean age was 53 (range 20 to 84) years. Most paid attention to medical statistics (6% paid no attention). The mean (SD) STAT-interest score was 68 (17) and ranged from 15 to 100. Confidence in using statistics was also high: the mean (SD) STAT-confidence score was 65 (19) and ranged from 11 to 100. STAT-interest and STAT-confidence scores were moderately correlated (r=.36, P<.001). Both scales demonstrated good test–retest repeatability (r=.60, .62, respectively), internal consistency reliability (Cronbach's α=0.70 and 0.78), and usability (individual item nonresponse ranged from 0% to 1.3%). Scale scores correlated only weakly with scores on a medical data interpretation test (r=.15 and .26, respectively).

CONCLUSION: The STAT-interest and STAT-confidence scales are usable and reliable. Interest and confidence were only weakly related to the ability to actually use data.

KEY WORDS: decision making; patient education; statistic.
DOI: 10.1111/j.1525-1497.2005.0179.x
J GEN INTERN MED 2005; 20:996–1000.

Patients are increasingly presented with medical statistics—that is, statements using numbers to describe the chance of various events (e.g., "FIBEREX™ lowers your cholesterol by 30%," "Mammograms lower a woman's chance of dying from breast cancer by a third," "Smokers are 10 times more likely to develop lung cancer"). Both the National Cholesterol Education Programs heart risk calculator[1] and the National Cancer Institute's Breast Cancer Risk Assessment Tool[2] provide users with quantitative estimates of the chance of heart disease or breast cancer (e.g., "your chance of breast cancer in the next 10 years is 0.4%"). Nearly all decision aids quantify the benefits and harms of different treatment options.

Nonetheless, it is not known whether patients want to see such statistics, whether they feel confident that they can interpret them, or whether their confidence is justified. The foregoing issues are important to any one designing patient communications. If statistics are not wanted—or if patients do not feel competent interpreting them—readers may ignore (or be intimidated by) materials with data. On the other hand, patients who want and expect data may be put off by materials that shun statistics and appear oversimplified. Finally, the extent to which confidence relates to ability is crucial. If confidence relates to ability, communicators could encourage those with too little confidence, and caution those with too much.

Existing measures of interest in and confidence interpreting medical statistics are limited. Interest measures are general (e.g., "do you want health information"),[3,4] and measures of confidence have only been used in the educational setting (e.g., asking college students whether the thought of taking another statistics course makes them "feel sick").[5–8] We developed and evaluated 2 new measures to assess attitudes toward medical statistics: interest in knowing them (STAT-interest) and confidence in one's ability to understand them (STAT-confidence). In addition, we examined relationships among an individual's interest, confidence, and ability to interpret medical statistics.

METHODS

Scale Development

The STAT-interest and STAT-confidence scales were developed as part of a larger project to teach people how to interpret medical statistics. We developed the interest and confidence items based on our own experience, medical literature reviews, and advice from experts in statistics, cognitive psychology, and education at Dartmouth College and Carnegie Mellon University. We revised items based on this feedback. All items underwent extensive pilot testing with patients and members of the general public (here the focus was on wording and understandability). Five items assessed interest; 3 items assessed confidence. According to the Flesch–Kincaid scoring algorithm (Microsoft Word X-Mac), the items are written at the eighth grade level.

The authors have no conflicts of interest to declare for this article or this research.

Dr. Schwartz is supported by Veterans Affairs Career Development Awards in Health Services Research and Development. Drs. Schwartz and Woloshin are supported by Robert Wood Johnson Generalist Faculty Scholar Awards and National Cancer Institute grants #CA91052-01 and CA104721. All the authors are supported by a Research Enhancement Award from the Department of Veterans Affairs. The views expressed herein do not necessarily represent the views of the Department of Veterans Affairs or the United States Government.

Address correspondence and requests for reprints to Dr. Woloshin: VA Outcomes Group (111B), Department of Veterans Affairs Medical Center, White River Junction, VT 05009 (e-mail: steven.woloshin@dartmouth.edu).

See editorial by Montori, p. 1071

Received for publication April 5, 2005
and in revised form April 8, 2005
Accepted for publication April 12, 2005

Scale Evaluation

Subjects. We recruited 224 English-speaking people from advertisements in local newspapers (n=104), the White River Junction VA outpatient clinic waiting area (n=76), and a hospital open house at Dartmouth Hitchcock Medical Center (n=44). The study was approved by the Dartmouth College IRB. Patient recruitment materials asked adults age 20 years and older to help us learn how to give people the information they need to make good medical decisions. We asked people to call a research assistant (who confirmed the subject's age). Participants were mailed a survey with a stamped return envelope. About 7 to 10 days after receipt of their completed survey, we mailed out a retest survey. Two hundred and twenty individuals (98%) completed the retest survey. Participants were paid $25.

Item and Scale Scores. For each item, we calculated the proportion of responses left blank to measure question usability. To create scores for each scale, we summed the points assigned to each response category. Most questions offered 5-level responses ranging from "strongly disagree" (assigned 0 points) to "strongly agree" (4 points) with a neutral middle category. We reversed the response order for negatively worded questions. Two questions used different response sets (3 and 4 levels). We revised the scoring for these questions so that all questions would contribute equally to the aggregate scores. We assigned the lowest response category "0" points, the highest category "4" points, and calculated intermediate values for the middle categories using a simple linear transformation (i.e., 3-level values were 0, 2, 4, and 4-level values were 0, 1.33, 2.66, 4). Aggregate scores were transformed onto a 0 to 100 point scale, with higher scores meaning more of the attribute (i.e., greater interest, greater confidence).

Test–Retest Repeatability. We calculated Pearson correlation coefficients of the test and retest scores for the 220 people who completed the retest survey 2 weeks later.

Reliability. We measured the internal consistency reliability with Cronbach's α, a measure of how responses to each question correlate with responses to all other questions in the scale.

Content Validity. We used feedback of experts in education, statistics, and cognitive psychology (locally and at Carnegie Mellon University) to assess the content validity of the measures, the extent to which the items capture the concept being measured without missing important aspects of the concept.

Construct Validity. The extent to which a measure actually captures the intended concept or construct was evaluated by analyzing responses to 2 additional questions in our survey meant to capture interest in statistics. The exact questions were:

"Next time I have to make a decision about my health, I plan to ....
a. Ask my doctor for statistics (yes/no)
b. Track down medical statistics (yes/no)

We hypothesized that respondents with greater interest in statistics would be more likely to actively try to find statistics when making medical decisions.

We also sought to learn whether people with greater interest in medical statistics or great confidence in interpreting them were better able to interpret medical statistics. Our measure of ability was the medical data interpretation test.[9] The 18-item data interpretation test measures the ability to make sense of and compare medical statistics about disease risk and risk reduction. Scores from this test range from 0 to 100 where higher scores represent higher abilities. The data interpretation test was modified early in the study; analyses relating attitudes to ability include the 175 participants who received the final version of the test.

Analyses

Means, standard deviations, and frequencies were used to describe item and scale scores. All comparisons were 2 sided and were considered statistically significant at $P<.05$. We used Pearson correlation to assess the association between interest, confidence, and medical data interpretation skills, analysis of variance to compare scores in different categories of educational attainment, and χ^2 for the construct validity analysis of interest score quartiles. We used STATA 8.0 (College Station, Tex).

RESULTS

The 224 participants represented a range of ages, incomes, and formal education although 52% had a college degree (Table 1). The mean age was 53 (range 20 to 84) years, 48% were female, and most were white. Tables 2 and 3 show the basic attributes of the individual items and the aggregate scale scores. Item nonresponse was low (ranging from 0% to 1.3%) for the interest and confidence questions.

Interest in medical statistics was fairly high: 94% paid some attention to medical statistics (36% paid considerable attention, and 58% a little) and 80% agreed that "to make wise decisions about my health, it is important to know how to interpret statistics." The STAT-interest score was normally distributed with a slight left skew; the mean score was 68 (SD 17), the median was 70, and scores ranged from 15 to 100. Interest in medical statistics was high at all education levels: mean scores were 60 for those with less than a high school education, 67 for high school graduates, 71 for college graduates, and 70 for postgraduates (P=.09).

Confidence in using statistics was also high: 72% thought statistics were easy to understand, and 79% felt "confident that I can make sense of medical statistics." The STAT-confidence scores were normally distributed with a slight left skew;

Table 1. Characteristics of the Study Sample (n=224)*

Mean age (range)	53 (20, 84)
Sex (% women)	48
Race (% white)	95
Household income (%)	
< $10,000	8
$10,000 to 24,999	22
$25,000 to 49,999	39
$50,000 to 99,999	22
≥ $100,000	9
Highest level of education (%)	
< High school graduate	6
High school degree	19
Some college	24
College degree	24
Postgraduate degree	28

Item nonresponse was: 10 (income), 3 (education).

Table 2. Responses to the STAT-Interest and STAT-Confidence Scales and their Items (n=224)*

Survey Introduction: *"We would like to learn how you feel about medical statistics. Here are some examples of what we mean by medical statistics:*
"FIBEREX lowers your cholesterol by 30%"
"Mammograms lower a woman's chance of dying from breast cancer by a third"
"Smokers are 10 times more likely to develop lung cancer"

STAT-interest items

	Yes, a lot (%)	*Yes, a little* (%)	*No, skip over* (%)
Do you usually pay attention to medical statistics in newspapers, magazines, or TV reports?	36	58	6

	Strongly disagree (%)	*Disagree* (%)	*Neither* (%)	*Agree* (%)	*Strongly agree* (%)
To make wise decisions about my health it is important to know how to interpret statistics.	1	4	15	47	33
I want my doctor to give me statistics when explaining tests or treatments.	1	6	15	52	26
When I am making a decision, I ask my doctor for medical statistics.	4	22	27	33	14
I do not believe in statistics because something will either happen or not happen to me.	20	40	29	11	1

STAT-confidence items

	Very easy (%)	*Easy* (%)	*Hard* (%)	*Very Hard* (%)
In general, how easy or hard do you find it to understand medical statistics?	11	61	27	1

	Strongly disagree (%)	*Disagree* (%)	*Neither* (%)	*Agree* (%)	*Strongly agree* (%)
I am confident that I can make sense of medical statistics.	1	12	18	57	12
I feel like I do not know how to interpret medical statistics.	16	42	25	13	4

the mean score was 64 (SD 19), the median was 65, and scores ranged from 11 to 100. Confidence in using medical statistics increased with educational attainment: mean scores were 56 (less than high school), 58 (high school graduate), 70 (college graduate), and 68 (postgraduate), $P<.001$.

As hypothesized, respondents with greater interest in statistics were more likely to say they would ask their doctor for or try to track down statistics the next time they had to make a decision about their health (Table 3): 46% of respondents with STAT-interest scores in the lowest quartile said they would ask their doctor for statistics and 19% said they would try to track down statistics; in contrast, the corresponding numbers for respondents with STAT-interest scores in the highest quartile were 97% and 80%, respectively ($P<.001$, both comparisons).

STAT-interest and STAT-confidence scores correlated only moderately with each other ($r=.36$, $P<.001$), suggesting that these are distinct constructs. Interest and confidence scores were only weakly correlated with performance on the medical data interpretation test ($r=.26$ ($P=.006$), $r=.15$ ($P=.04$)).

DISCUSSION

Proponents of shared decision making often assume that patients desire to have statistics on the benefits and harms of medical treatments.[11–15] Nonetheless, the extent to which patients actually want data—or feel confident in interpreting them—is unknown, prompting us to develop the new measures. Such measures are important to help communicators design materials, and to help judge their impact. Knowing how people feel about data could lead to tailored messages that might be more effective. Such measures would also be useful in the evaluation of educational interventions designed to promote better data interpretation skills.

We developed 2 new measures for assessing attitudes toward medical statistics. Both measures demonstrated good psychometric properties: very low item nonresponse; broad response ranges; and substantial test–retest and internal consistency reliability. It should be noted that while Cronbach's α was "respectable" for both scales,[10] a higher internal consistency would be desirable if the scales were used to track individuals rather than groups.

Three findings emerged from our study: scores on both the interest in and confidence in interpreting measures were generally high; interest and confidence are related but distinct concepts; and neither implies ability.

Our findings need to be interpreted in the light of several limitations. First, some may disagree with the content of the measures. We selected items based on our experience, a literature review, and discussions with experts in education, statistics, and psychology. While this study suggests that our questions work well, others may have chosen different items or wordings. Second, the study participants were a convenience sample of individuals recruited with an advertisement reading "help us learn how to provide medical information to promote wise decision making." Even though we did not specifically recruit people interested in "statistics," people interested in medical information may be different from the general population. This makes the (limited) observed relationships

Table 3. Psychometric Properties for the Attitudes Toward Medical Statistics Scale

Property	Measure Used	STAT-Interest	STAT-Confidence
Item nonresponse	Proportion not answering question Mean (range)	0.08% (0% to 0.4%)	0.4% (0% to 1.3%)
Score distribution	Mean score (standard deviation), median and range	68 (SD=17) Median 70 Range 15 to 100	65 (SD=19) Median 65 Range 11 to 100
Test–retest repeatability	Correlation of scores at test and retest 2 wk later (n=220)	Pearson r=.60	Pearson r=.62
Internal consistency reliability	Cronbach's α (goal is 0.7 to 0.8; lower value suggests lack of coherence as a scale, and higher value suggests that items are redundant)	α=0.70	α=0.78
Content validity	Initial questionnaires revised based on feedback from experts in education, statistics, and cognitive psychology		
Construct validity (interest scale only)		*"Next time I have to make a decision about my health, I plan to..." (%yes)*	
Interest in statistics score		Ask my doctor for statistics (%)*	Track down statistics (%)*
Lowest quartile (15 to 54)		46	19
2nd quartile (55 to 69)		81	49
3rd quartile (70 to 79)		96	68
Highest quartile (80 to 100)		97	80

*P<.001 (χ^2).
r, *correlation coefficient.*

between our 2 measures and ability all the more striking. Third, further study among those with very little formal education will be important. Although usability was very good among participants with less than a college degree (almost all of these respondents were able to complete every question), our sample included few respondents with less than a high school education.

Finally, some may be concerned that the confidence was only weakly related to ability to use data. Many may assume that confidence ought to predict ability. We hypothesized that confidence using statistics would be positively correlated with ability, but did not expect the correlation to be strong as people with considerable ability might realize that things can be very complicated (and feel daunted), while people with little ability might not realize there is a lot they do not know (and feel inappropriately confident). In our literature review, we could not find an existing, validated "confidence using statistics" measure. We did find a few published studies exploring 2 related concepts: attitudes toward studying statistics, and math anxiety.[5–8] Our hypothesis about how confidence would relate to ability is supported by these studies: in each case investigators found a weak to moderate positive relationship between the "confidence"—type measure (e.g., "the thought of taking another statistics course makes me feel sick") and the ability measure (e.g., statistics course grade).

That confidence was only weakly related to the ability to use data should serve as a warning to communicators: people may overestimate their ability to use statistics. Just because someone expresses interest in viewing data (efforts to let patients select among a variety of presentation formats are increasingly common), their ability to make sense of the information should not be assumed; instead, comprehension testing should probably be built in to the materials. The best data in the most preferred format are still useless if people do not understand them.

The authors would like to thank Baruch Fischhoff, PhD, Carnegie Mellon University, and Laurie Snell, PhD. Dartmouth College, for their help in reviewing drafts of instruments.

REFERENCES

1. Risk assessment tool for estimating 10-y risk of developing hard CHD. Third report of the expert panel on detection, evaluation, and treatment of high blood cholesterol in adults. National cholesterol education program. http://hin.nhlbi.nih.gov/atpiii/calculator.asp?usertype=prof. Accessed August 12, 2004.
2. Breast cancer risk assessment tool. National Cancer Institute. http://bcra.nci.nih.gov/brc/. Accessed August 12, 2004.
3. **Nease RF, Brooks WB.** Patient desire for information and decision making in health care decisions: the autonomy preference index and the health opinion survey. J Gen Intern Med. 1995;10:593–600.
4. **Ende J, Kazis L, Ash A, Moskowitz M.** Measuring patients' desire for autonomy. J Gen Intern Med. 1989;4:23–30.
5. **Roberts D, Bilderback E.** Reliability and validity of a statistics attitude survey. Educ Psychol Meas. 1980;40:235–8.
6. **Roberts D, Saxe J.** Validity of a statistics attitude survey: a follow-up study. Educ Psychol Meas. 1982;42:907–12.
7. **Schultz K, Koshino H.** Evidence of reliability and validity for wise's attitude toward statistics scale. Psych Rep. 1998;82:27–31.
8. **Suinn R, Winston E.** The mathematics anxiety rating scale, a brief version: psychometric data. Psych Rep. 2003;92:167–73.
9. **Woloshin S, Schwartz LM, Welch HG.** Assessing patient's ability to interpret medical data: a validation study. Med Dec Making. 2005;25:290–300.
10. **DeVellis RF.** Scale Development: Theory and Applications. 2nd ed. Thousand Oaks, Calif: Sage Publications; 2003.
11. **Department of Veterans Affairs.** VHA Notice 99–02. Shared decision making. June 15, 1999. http://www.va.gov/vhapublications/ViewPublication.asp?pub_ID=1081. Accessed April 8, 2004.

12. Health Research: What's in it for consumers? Report of the Standing Advisory Group on Consumer Involvement in the NHS Research and Development Programme. 1997. www.invo.org.uk/pdf/whats_in_it_for_consumers.pdf. Accessed August 12, 2004.
13. **Committee on Quality of Health Care in America. Institute of Medicine.** Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.
14. **Sheridan S, Harris R, Woolf S.** 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.
15. **Winker MA, Flanagin A, Chi-Lum B, et al.** Guidelines for medical and health information sites on the internet: principles governing AMA web sites. American Medical Association. JAMA. 2000;283:1600–6.