A Randomized Comparison of Patients' Understanding of Number Needed to Treat and Other Common Risk Reduction Formats

Stacey L. Sheridan, MD, MPH, Michael P. Pignone, MD, MPH, Carmen L. Lewis, MD, MPH

BACKGROUND: Commentators have suggested that patients may understand quantitative information about treatment benefits better when they are presented as numbers needed to treat (NNT) rather than as absolute or relative risk reductions.

OBJECTIVE: To determine whether NNT helps patients interpret treatment benefits better than absolute risk reduction (ARR), relative risk reduction (RRR), or a combination of all three of these risk reduction presentations (COMBO).

DESIGN: Randomized cross-sectional survey.

SETTING: University internal medicine clinic.

PATIENTS: Three hundred fifty-seven men and women, ages 50 to 80, who presented for health care.

INTERVENTIONS: Subjects were given written information about the baseline risk of a hypothetical "disease Y" and were asked (1) to compare the benefits of two drug treatments for disease Y, stating which provided more benefit; and (2) to calculate the effect of one of those drug treatments on a given baseline risk of disease. Risk information was presented to each subject in one of four randomly allocated risk formats: NNT, ARR, RRR, or COMBO.

MAIN RESULTS: When asked to state which of two treatments provided more benefit, subjects who received the RRR format responded correctly most often (60% correct vs 43% for COMBO, 42% for ARR, and 30% for NNT, P = .001). Most subjects were unable to calculate the effect of drug treatment on the given baseline risk of disease, although subjects receiving the RRR and ARR formats responded correctly more often (21% and 17% compared to 7% for COMBO and 6% for NNT, P = .004).

CONCLUSION: Patients are best able to interpret the benefits of treatment when they are presented in an RRR format with a given baseline risk of disease. ARR also is easily interpreted. NNT is often misinterpreted by patients and should not be used alone to communicate risk to patients.

KEY WORDS: data interpretation (statistical); decision making; numeracy; patient participation (statistics and numerical data).
J GEN INTERN MED 2003; 18:884–892.

In a recent report, the Institute of Medicine recognized "patient centeredness" as a key component of health care quality, stating that a provider-patient partnership is needed "to ensure that decisions respect patients' wants, needs, and preferences and that patients have the education and support they require to make decisions and participate in their own care."[1] With this emphasis on involving patients in decision making, clinicians are increasingly challenged to communicate health information to patients in unbiased and easily understandable ways. This is especially true for decisions in which the balance of potential harms and benefits is a close call. In those decisions, patients must understand quantitative information to compare the potential harms and benefits and choose the health alternative that is most consistent with their values. For clinicians, this raises questions about how to communicate a clear picture of the harms and benefits associated with each decision.

For communicating treatment benefits, some have suggested that quantitative information may be best interpreted when it is presented as number needed to treat (NNT).[2,3] Number needed to treat is an empirically derived estimate of the number of patients who must be treated in order to expect that one patient will avoid an adverse event or outcome over a defined period of time. Mathematically, NNT is the reciprocal of the absolute risk reduction (ARR: the decrease in disease incidence due to treatment), and therefore provides an estimate of absolute patient benefit. Enthusiasm for its use arose because of its apparent utility for comparisons of treatments, and harms and benefits,[4,5] and because of patients' difficulties in understanding other risk reduction formats.[3] Criticism against its use has centered around its statistical properties[6] and its name, which encourages individuals to think of it as precise and without probabilistic content.[6]

Studies examining the effects of presenting treatment benefit information to patients in alternate risk reduction formats have examined the persuasiveness of NNT, not a patient's ability to understand it. Hux and Naylor[7] have compared the willingness of patients to accept a treatment whose benefits were presented alternately as NNT, ARR, or relative risk reduction (RRR: the decrease in disease incidence relative to those who are not taking treatment). They found that patients were more willing to accept a treatment when benefits were presented as an RRR rather than as an ARR or an NNT. Such studies may indicate that NNT is less persuasive to patients than other presentations of treatment benefits. These studies, however, did not examine a patient's ability to correctly interpret the information they received; willingness to accept treatment is dependent on patient values, not solely on correct interpretation of information

on treatment benefits. At present, we know of no empiric evidence about whether patients understand information on treatment benefits better when they are presented as NNT compared to other common risk formats.

Our study examined whether patients better understand written information on treatment benefits when it is presented as NNT, ARR, RRR or COMBO. Because understanding is a complex process that is not easily measured, we used patient ability to correctly interpret information on treatment benefits as a reasonable and close approximation for understanding.

METHODS

Study Participants

After approval from our university institutional review board, we surveyed men and women, ages 50 to 80, who presented for care at a university internal medicine clinic. Patients were excluded if this was their first visit to the clinic, if they reported that they were unable to understand, speak, or read English, or if they had previously participated in the survey. Potential participants were identified from daily clinician schedules and were approached about the study in the clinic waiting room or in their exam room while they waited for the clinician. When our research assistant was not interviewing other study patients, she approached every next eligible and available patient who was between the ages of 50 and 80 and presented to our clinic for a return visit. Overall, we estimate that she approached 60% of patients ages 50 to 80 who presented to our clinic for a return visit between June and November 2000.

Questionnaire Design

Presentation of Information on Treatment Benefit. Written information about the baseline risk of a hypothetical disease Y and the benefits of two hypothetical treatments was presented to each subject in one of four risk reduction formats: NNT, ARR, RRR, or a combination of all three of these formats (see Fig. 1). Questionnaires differed only in this risk presentation.

Assessment of Ability to Interpret Treatment Benefit. Subjects were given information about the baseline risk of a hypothetical disease Y and were asked (1) to state which of two drug treatments for disease Y provided more benefit, and (2) to calculate the effect of one of these drug treatments on the given baseline risk of disease:

(1) Compare treatments A and B. Which treatment is more effective?

     1. Don't know
     2. A is more effective than B
     3. B is more effective than A
     4. A and B are equally effective

(2) What is the chance that you will develop disease Y after treatment A? ___ out of 1,000

Responses to these two tasks were counted as either correct or incorrect according to the answers shown in Figure 1.

Assessment of Numeracy. We assessed numeracy using the three-question numeracy scale shown in Figure 2.[3] Answers to these questions were counted as either correct or incorrect according to the calculations that are shown.

Assessment of Prior Exposure to Risk Concepts. To assess patients' familiarity with risk concepts, we measured several indicators of prior exposure to risk concepts: education, including statistics or epidemiology training; discussion of a medical decision with a physician; prior quantitative risk discussions with a physician; and self-perceived facility with numbers.

Assessment of Demographic Information. The demographic characteristics of each study participant were assessed with single-item questions on age, race/ethnicity, education, and self-perceived health status.

Randomization

We used a computerized random number generator to assign risk presentation formats to each consecutively recruited subject. Format assignments were sealed in security envelopes until just prior to questionnaire administration. After a subject agreed to participate in the survey, the research assistant broke the seal on the security envelope, determined the assignment, and gave the subject the appropriate version of the questionnaire.
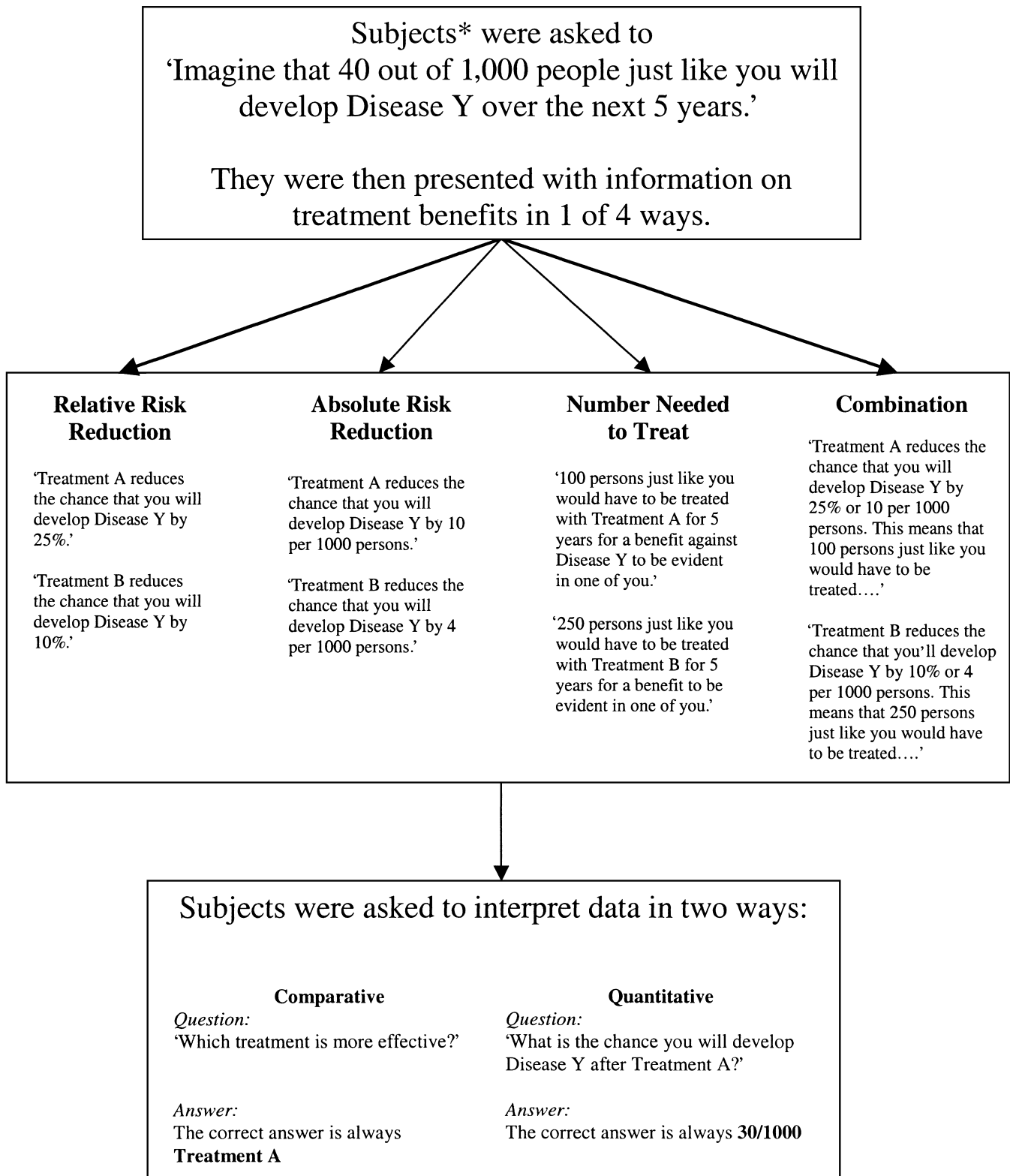
Questionnaire Administration

Questionnaires were self-administered and took approximately 10 minutes to complete. We did not allow subjects to ask the research assistant for clarification on any questions; we did, however, perform cognitive and pilot testing of the questionnaire on both first-year medical students and patients prior to survey administration to minimize unanswered questions due to survey burden or confusion over question wording.

We collected all questionnaires before subjects left the clinic. Questionnaires were counted as complete and were included in the analysis if they had any marks on the last two pages, which included three questions on numeracy and two questions on treatment benefit.

Sample Size Calculation

We determined that a sample size of 100 in each risk reduction format group was required to detect a 15% difference in the proportion of people correctly comparing and calculating the effects of treatment on a given baseline risk of disease, accepting an alpha of 0.05 and a power of 80%. We considered 15% a clinically important difference. Additionally, a recent study had shown an approximately 15% difference in the proportion of women who were able to correctly apply information about mammography benefits when the information was presented alternately as ARR or RRR in the context of baseline risk.[3]

Subjects* were asked to
'Imagine that 40 out of 1,000 people just like you will develop Disease Y over the next 5 years.'

They were then presented with information on treatment benefits in 1 of 4 ways.

Relative Risk Reduction

'Treatment A reduces the chance that you will develop Disease Y by 25%.'

'Treatment B reduces the chance that you will develop Disease Y by 10%.'

Absolute Risk Reduction

'Treatment A reduces the chance that you will develop Disease Y by 10 per 1000 persons.'

'Treatment B reduces the chance that you will develop Disease Y by 4 per 1000 persons.'

Number Needed to Treat

'100 persons just like you would have to be treated with Treatment A for 5 years for a benefit against Disease Y to be evident in one of you.'

'250 persons just like you would have to be treated with Treatment B for 5 years for a benefit to be evident in one of you.'

Combination

'Treatment A reduces the chance that you will develop Disease Y by 25% or 10 per 1000 persons. This means that 100 persons just like you would have to be treated....'

'Treatment B reduces the chance that you'll develop Disease Y by 10% or 4 per 1000 persons. This means that 250 persons just like you would have to be treated....'

Subjects were asked to interpret data in two ways:

Comparative
Question:
'Which treatment is more effective?'

Answer:
The correct answer is always **Treatment A**

Quantitative
Question:
'What is the chance you will develop Disease Y after Treatment A?'

Answer:
The correct answer is always **30/1000**

* A short introduction told subjects that information was about two treatments for a single disease and that there was no difference between the treatments in cost, number of medication doses to be taken per day, or side effects.

FIGURE 1. Presentation of treatment benefits and assessment of subject ability to correctly interpret them.

Each subject answered three questions to assess their skill with numbers:

	Question	Calculation	Correct Answer
(1)	Imagine that we flip a fair coin 1000 times. What is your best guess about how many times the coin would come up heads?	½ (1000)	500
(2)	In the lottery, the chance of winning a prize is 1%. What is your best guess about how many people would win a prize if 1000 people each buy a single ticket to the lottery?	1/100 = X/1000	10/1000
(3)	In the publishing sweepstakes, the chance of winning a car is 1 in 1000. What percent of tickets to the publishing sweepstakes wins a car?	1/1000 = X/100	0.1%

FIGURE 2. Assessment of numeracy.

Statistical Analysis

To assess the success of randomization, we compared the characteristics of subjects who received each risk presentation format, using χ^2 tests for categorical variables and t tests for continuous variables. We also used χ^2 tests and t tests to examine the relationships between risk reduction format and (1) the ability to correctly assess which of two treatment benefits provided greater benefit, and (2) the ability to correctly calculate the effect of treatment on a given baseline risk of disease. Fisher's exact tests were used when comparisons involved a small number of subjects. The relationships between each baseline characteristic and subject ability to correctly perceive treatment benefit was calculated in a similar manner. Due to multiple testing, we considered an alpha of 0.01 to be statistically significant.

RESULTS

A total of 623 patients were approached to participate in our survey (see Fig. 3). Fourteen percent were ineligible and 22% refused. Three hundred ninety-eight eligible patients returned their questionnaires for a response rate of 74%. Ninety percent of returned questionnaires were complete according to our criteria, and 94% of these had pen marks indicating consideration of at least 1 of the 2 questions regarding treatment benefit. Ninety percent had an answer for the comparison of the benefits of 2 treatments and 65% had an answer for the effect of treatment on a given baseline risk of disease; in the remaining 10% and 35% of questionnaires in which no answer was provided, the blanks were counted as incorrect answers.

Table 1 provides information about the study participants. Subjects in the NNT group were somewhat more likely to be male and white. Accounting for multiple comparisons, however, the baseline characteristics of the four groups were statistically similar.

Ability to Interpret Treatment Benefit

Subjects correctly identified which of 2 treatments provided greater benefit 44% of the time, but correctly calculated the exact effect of treatment benefit on a given risk of disease only 13% of the time. The format in which information on treatment benefits was presented had a strong effect on subjects' ability to compare and calculate treatment benefits (see Fig. 4).

When comparing 2 treatments, 30% of subjects who received information as NNT correctly stated which treatment provided more benefit compared with 60% of subjects who received RRR, 42% who received ARR, and 43% who received the COMBO ($P = .001$). When calculating the effect of treatment on a given baseline risk of disease, however, only 6% of subjects who received information as NNT correctly stated which treatment provided more benefit, compared with 21% who received RRR, 17% who received ARR, and 7% who received the COMBO ($P = .004$).

In the NNT group, 39% of subjects provided no answer when asked to calculate the exact effect of treatment on

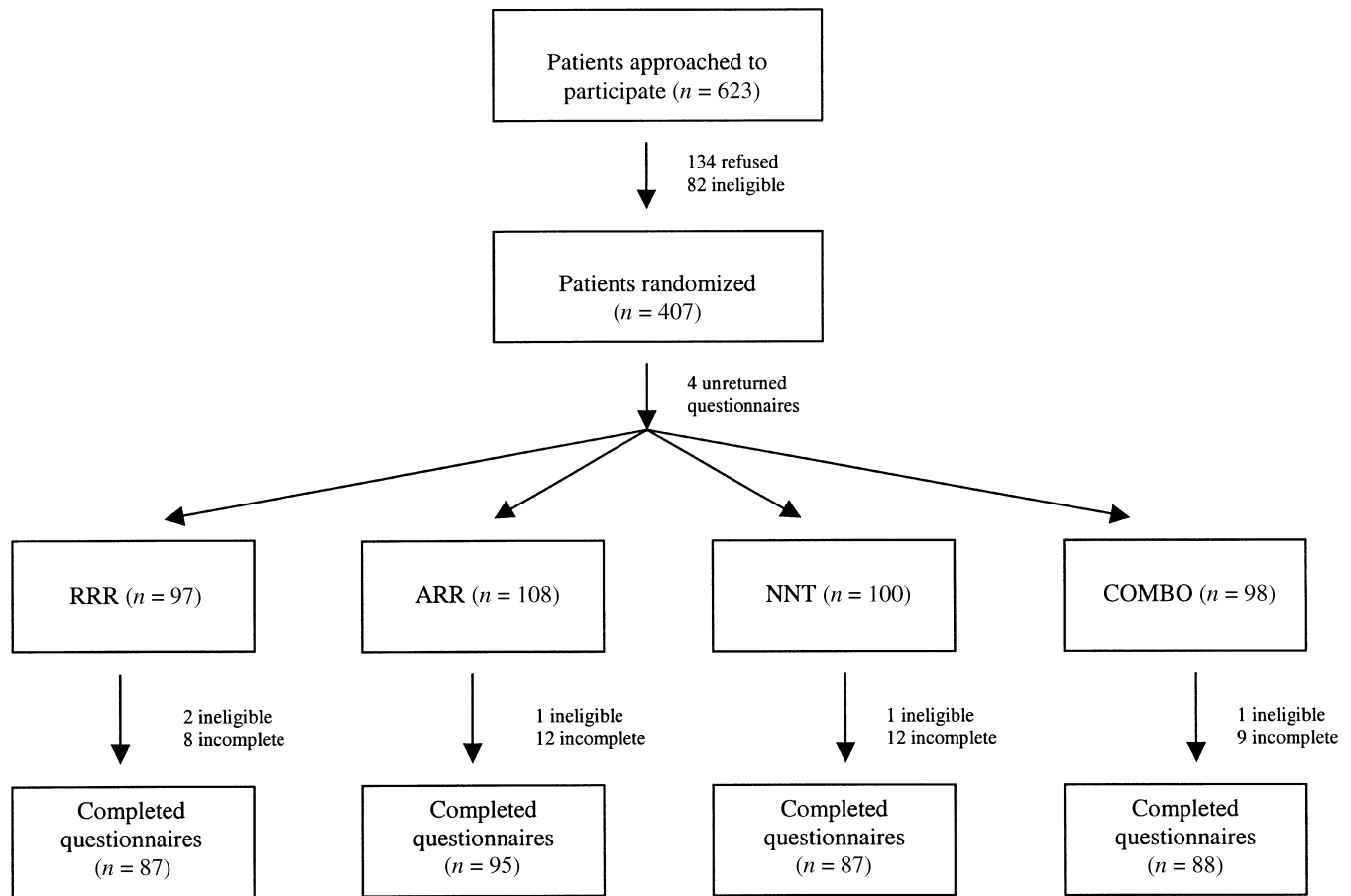**FIGURE 3.** Patient accrual.

a given baseline risk of disease, compared with 26% of subjects in the RRR group, 32% in the ARR group, and 42% in the COMBO group ($P = .12$). Of those who calculated the exact effect of treatment on the given baseline risk of disease, 15% were off by an order of magnitude or more: 25% in the NNT group, 11% in the RRR group, 17% in the ARR group, and 8% in the COMBO group ($P = .08$). Interestingly, a substantial portion of each group (25% in the NNT group, 19% in the RRR group, 38% in the ARR group, and 45% in the COMBO group; $P = .008$) reported that the correct

Table 1. Subject Characteristics by Risk Presentation Format

Characteristic	COMBO (N = 88)	RRR (N = 87)	ARR (N = 95)	NNT (N = 87)	ALL (N = 357)	P Value*
Mean age, y	63	62	62	64	63	.57
Male, %	32	35	27	48	35	.03
White, %	60	76	62	79	69	.01
Fair/poor health, %	50	49	42	56	51	.29
Discussion of medical decision with a doctor, %	56	72	59	63	62	.12
Receiving some quantitative information from a doctor, %	13	15	13	12	13	.50
With some college education, %	51	61	54	66	58	.52
With statistics training, %	16	19	17	26	20	.39
Consider themselves good with numbers, %	70	71	70	70	70	.99
Answering 3 numeracy questions correctly, %	0	2	3	3	2	.60
Answering 2 numeracy questions correctly, %	30	23	25	32	27	
Answering 1 numeracy question correctly, %	27	37	27	28	30	
Answering no numeracy questions correctly, %	43	38	44	37	41	

* P value for any difference in patient characteristics among the 4 risk presentation format groups.
COMBO, combination of all three risk reduction presentations; RRR, relative risk reduction; ARR, absolute risk reduction; NNT, number to treat.

A. Ability to Correctly Compare Treatment Benefits

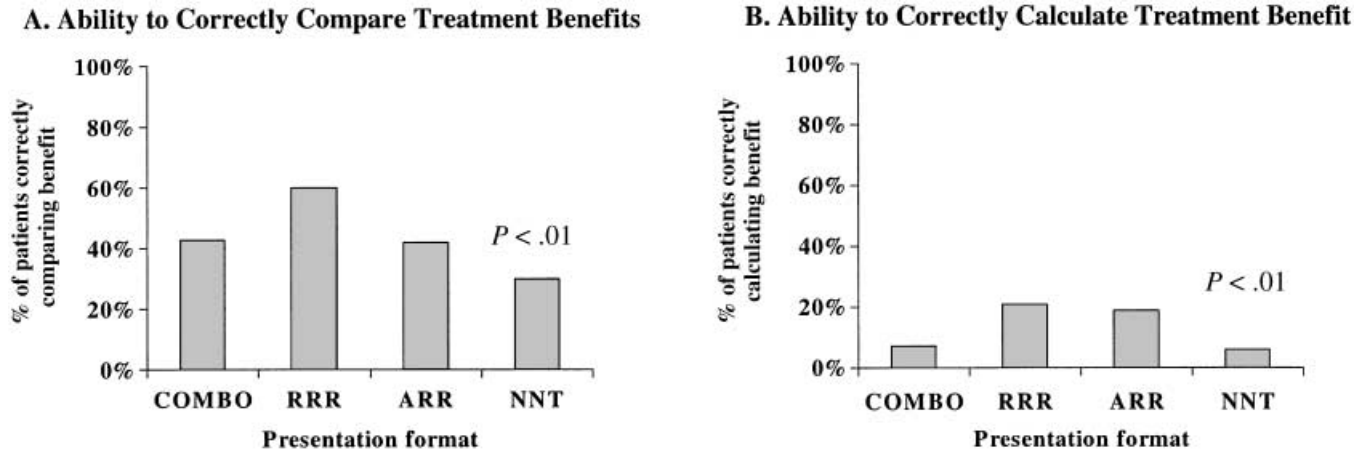## B. Ability to Correctly Calculate Treatment Benefit

FIGURE 4. Risk reduction format and the ability to correctly interpret treatment benefits.

answer was "10 per 1,000," which is the magnitude of the treatment benefit, not the risk of disease after treatment.

Numeracy

Although 70% of subjects perceived themselves to be good with numbers, only 2% of subjects answered all three numeracy questions correctly: 28% answered 2 numeracy questions correctly and 71% answered 1 (30%) or no (41%) numeracy questions correctly (see Table 1). As expected, a subject's level of education was correlated with their numeracy score (Pearson's correlation coefficient = 0.55). Nonetheless, 56% of subjects who had at least some college education answered 1 or no numeracy questions correctly.

Relationship Between Numeracy and the Ability to Interpret Treatment Benefit

Patients with better numeracy skills correctly compared and calculated treatment benefits more often (see Fig. 5). Eighty-eight percent of subjects who gave 3 correct answers to the numeracy questions correctly stated which treatment provided more benefit, whereas only 63% of

subjects who gave 2 correct answers to the numeracy questions, and 35% of subjects who gave 1 or no correct answers to the numeracy questions, correctly did so ($P < .001$). Similarly, 50% of subjects who gave 3 correct answers to the numeracy questions correctly calculated the effect of treatment on a given baseline risk of disease, compared with 30% of subjects who gave 2 correct answers to the numeracy questions, and 5% of subjects who gave 1 or no correct answers to the numeracy questions ($P < .001$).

Relationship Between Subjects' Demographics and the Ability to Interpret Treatment Benefit

Subjects' ability to compare and calculate treatment benefits was affected by their baseline characteristics (see Table 2). Subjects had more difficulty with both comparisons and calculations if they were female, nonwhite, had no college education, were in poor health, or had no prior quantitative discussions with their physicians.

DISCUSSION

Although some have suggested that quantitative treatment benefit information may be best interpreted when

A. Ability to Correctly Compare Treatment Benefit

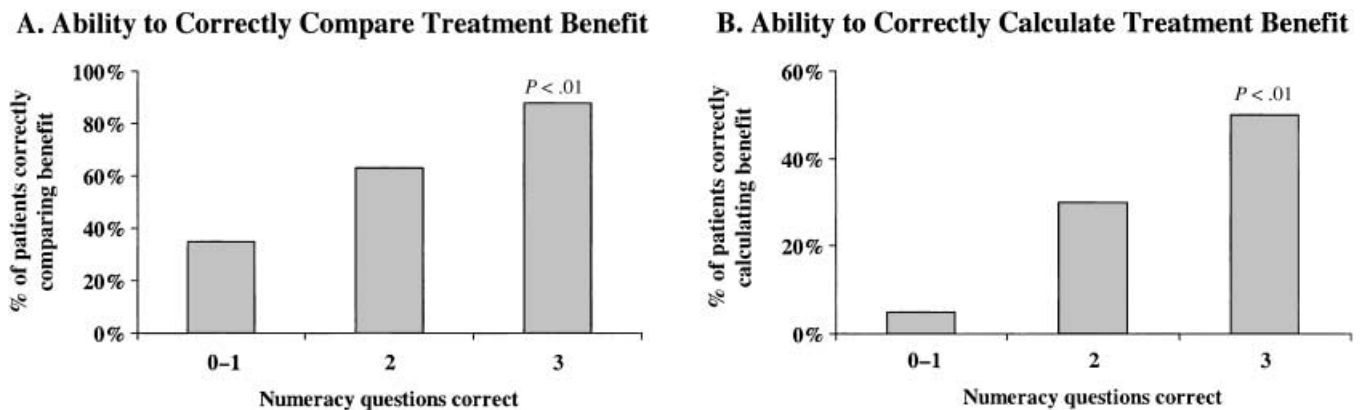## B. Ability to Correctly Calculate Treatment Benefit

FIGURE 5. Numeracy and the ability to correctly interpret treatment benefits.

Table 2. Percentage of Subjects Correctly Comparing and Calculating Treatment Benefits, According to Their Baseline Characteristics (N = 357)

Characteristic	N	% Correct Comparison*	P Value	% Correct Calculation*	P Value
Sex					
Male	126	54	.004	24	<.001
Female	229	38		7	
Race					
White	245	54	<.001	18	<.001
Nonwhite	110	20		1	
Education					
Some college	203	56	<.001	21	<.001
No college		41		2	
Health status					
Fair/poor	180	36	.002	4	<.001
Good/excellent	175	52		22	
Discussion of medical decision with an MD					
Yes	223	51	<.001	17	.001
No	134	31		5	
Information presentation by a doctor					
Some quantitative	41	76	<.001	37	<.001
Qualitative	216	44		13	
Consider themselves good with numbers					
Yes	234	50	.002	17	<.001
No	99	31		3	
Statistical training					
Some	68	66	<.001	35	<.001
None	280	39		7	

* % and P *values based on* χ^2 *tests or Fischer's exact test when group numbers <5.*

it is presented as NNT,[2,3] our study suggests just the opposite. Patients had more difficulty interpreting written treatment benefit information when it was presented as NNT. This effect was evident whether patients were comparing the benefits of two treatments or calculating the exact effect of a treatment on a given baseline risk of disease. The difficulty was magnified in patients with lower numeracy levels.

Patients' difficulties with NNT could perhaps have been predicted from the results of several studies examining the perception of risk in both rate (X in 1,000) and proportion (1 in X) formats.[8–10] All of these studies found that patients had more difficulty with the "1 in X" scale, perhaps because larger numbers are represented by smaller numbers in the denominator. Because NNT is essentially a "1 in X" scale, reporting the number of people who must be treated for 1 person to benefit, it is not surprising that patients would have difficulty comparing and calculating treatment benefits presented as NNT.

This study suggests that written information on treatment benefits is better understood when it is presented as ARR or RRR in the context of a given baseline risk of disease. When comparing the effectiveness of 2 treatments, both the ARR and RRR require equivalent and straightforward tasks: the patient must choose the treatment with the largest risk reduction. When calculating the effect of a given treatment on a baseline risk, the ARR requires the

simplest task: subtraction. The RRR presentation, however, requires a very familiar task, a task akin to figuring out how much money would be saved during a sale at the store. Familiarity may serve to smooth differences attributable to format. The effect of the combination presentation is more difficult to characterize. It is as easily interpretable as the ARR and RRR presentations when patients are asked to compare the effectiveness of 2 treatments, but performs no better than the NNT presentation when patients are asked to calculate the effect of a treatment on a given baseline risk of disease. This result is harder to explain, but may be due to overload of information or the more difficult construct of the presentation format. Regardless, the poor performance of patients presented with the combination presentation when they are trying to calculate the effect of treatment on a given risk of disease makes the presentation format a less desirable one for clinicians. More information is not necessarily better.

Because even patients who received the "simplest" risk presentation formats had difficulty comparing and calculating treatment benefit information, this study again raises questions about whether patients can independently make informed medical decisions using written quantitative information. We did note that patients who had a recent medical discussion with their doctor, or who reported receiving at least some quantitative information from their doctor,

interpreted treatment benefits correctly more often than patients who did not report these interactions. This finding may indicate that patients who are more educated, or have better numeracy skills, are more likely to receive quantitative information from their physicians. Alternately, it may indicate that patients can learn the skills needed to interpret quantitative presentations of treatment benefits through discussions with their doctor. We are aware of at least one effort to prepare patients to better interpret quantitative information on treatment benefit through a computerized risk tutorial.[11] Additionally, several researchers continue to explore more accessible ways to present quantitative information. A recent comparison of graphical and numerical presentations of treatment benefit, however, showed that numbers were interpreted equally well (in comparison tasks) or better (in calculation tasks) than graphical presentations.[12] Thus, a continued effort to improve patient interpretation of numerical treatment benefits is indicated.

The Evidence-Based Medicine Working Group has recently proposed presenting patients with the Likelihood of Being Helped Versus Harmed as a means of communicating treatment benefit.[13] This treatment benefit presentation format incorporates an individual patient's values with the number of patients that need to be treated for a benefit to be realized in 1 and the number of patients that need to be treated for a harm to be realized in 1, to report that a patient is x times more likely to be helped than harmed. The Likelihood of Being Helped Versus Harmed format avoids the problematic "1 in x" scale of the Number Needed to Treat and obviates the need for a patient to calculate the exact effect of a treatment on a given baseline risk of disease, because the goal of this calculation is to facilitate the weighing of harms and benefits. This presentation format therefore deserves further testing. Until this or the other innovations in the quantitative presentation of treatment benefits are tested, our study supports presenting treatment benefit information to patients as ARR or RRR, when a baseline risk of disease is available, and verifying patient understanding.

Our study does have several potential limitations. First, written information on treatment benefits was presented out of context in this study, reducing patients' personal involvement in the tasks measuring perception of treatment benefit. Previous research has shown that the degree of issue involvement influences how a patient processes information:[14] those who are highly involved process information in a detailed and integrative way, whereas those who are less involved process information superficially. Involving patients in actual treatment decisions would be expected to increase their processing of quantitative information, although these effects may be diminished by the burden of acute illness, which may make patients less able to process complex information. Regardless, we expect that the out of context presentation would affect all risk presentation format groups equally.

Second, we asked subjects to interpret individual benefit from NNT. This is not the task for which NNT was proposed. This is, however, the task implied by those who claim that NNT is easily understood by patients; patients are intrinsically interested in how an intervention affects them, not how an intervention affects the population from which probabilistic information was derived.

Third, ARR, RRR, and NNT can be worded in many different ways. Whether alternate wording of the presentation of treatment benefits would produce different results has not been tested. In our study, the readability of treatment benefits varied by risk reduction format (Fleisch-Kincade grade levels 5.8 [RRR], 8.3 [ARR], 11.5 [NNT], and 10.9 [COMBO], despite the fact that the readability of the entire presentation regarding treatment benefits was similar (Fleisch-Kincade grade levels from 10.3 to 11.8). It may be possible to word treatment benefit presentations so that they are less different in reading grade levels. Future research will help us determine what proportion of the difference in patient understanding is from differences in the readability of the presentations versus differences inherent to the concepts themselves.

Fourth, we did not measure literacy. Inability to understand the written presentations of treatment benefit and the written questions (Fleisch-Kincade grade level 8) could have accounted for some of our findings. Whether presenting the information orally would change results should be investigated in future studies.

Fifth, subjects in this study had no opportunity to ask questions about treatment benefits. It is possible simple clarification from a physician may have significantly improved patient understanding of some of these risk reduction formats.

Sixth, we used between-subject comparisons rather than within-subject comparisons of the risk reduction formats to reduce the length of the survey, increase the feasibility of survey administration in our clinic setting, and reduce "training" effects. Although within-subject comparisons have the advantage of allowing each subject to act as his own control, we were able to study large numbers of patients to minimize the effect of between-subject variation. Additionally, the findings of three within-group comparisons[15–17] and three between-group comparisons[18–20] of the persuasiveness of alternate risk reduction formats in physicians have showed a high degree of consistency.

Seventh, our results may not be generalizable to patients in other age groups. Older adults are more likely than middle-aged and younger adults to demonstrate limited numeracy skills.[21]

Finally, the nonconsecutive nature of our sample may also affect generalizability. When two eligible patients presented to the clinic at the same time, our sole research assistant could approach only one. We are not aware of appreciable differences between the patients who were approached to participate in the study and those who were not, but we made no formal attempt to monitor differences. Similarly, we do not have information on the patients who refused to participate in our study. We suspect, however, that people who did not participate were less confident in their quantitative abilities than those who participated.

Despite these limitations, this study provides important information to clinicians who wish to help their patients make informed decisions: many patients have poor numeracy skills; patients have difficulty interpreting quantitative information; and NNT, and sometimes combination presentations, are interpreted less successfully than ARR or RRR.

To address patients' limited ability to use quantitative information, clinicians may, in the short term, want to use written quantitative information only with patients with higher numeracy skills; present information that uses comparison, not calculation; present risk reduction information as ARR or RRR rather than as NNT or a combination presentation; and verify patient understanding after the presentation of treatment benefit information.

In the longer term, however, we believe researchers should continue to explore the robustness of current observations on ARR, RRR, and NNT in different populations and in different risk reduction scenarios, with alternate wording and different combinations of ARR, RRR, and NNT. Both clinicians and researchers should also work to improve patient understanding of quantitative information through exploring new presentation formats and developing patient tutorials on how to interpret quantitative health information.

REFERENCES

1. Institute of Medicine (US). Envisioning the National Health Care Quality Report. Washington, DC: The Institute; 2001.
2. Rajkumar S, Sampathkumar P, Gustafson A. Number needed to treat is a simple measure of treatment efficacy for clinicians. J Gen Intern Med. 1996;11:357–9.
3. 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.
4. Laupacis A, Sackett D, Roberts R. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med. 1988;318:1728–33.
5. Sackett D. Inference and decision at the bedside. J Clin Epidemiol. 1989;42:309–16.
6. Hutton JL. Number needed to treat: properties and problems. J Royal Statist Soc A. 2000;163:403–19.
7. Hux JE, Naylor CD. Communicating the benefits of chronic preventive therapy: does the format of efficacy data determine patients' acceptance of treatment? Med Decis Making. 1995;15:152–7.
8. Grimes DA, Snively GR. Patients' understanding of medical risks: implications for genetic counseling. Obstet Gynecol. 1999;93:910–4.
9. Woloshin S, Schwartz L, Byram S, Fischhoff B, Welch H. A new scale for assessing perceptions of chance: a validation study. Med Decis Making. 2000;20:298–307.
10. Lipkus I, Samsa G, Rimer B. General performance on a numeracy scale among highly educated samples. Med Decis Making. 2001;21: 37–44.
11. Woloshin S, Schwartz LM. How can we help people make sense of medical data? Eff Clin Prac. 1999;2:176–83.
12. Feldman-Stewart D, Kocovski N, McConnell BA, Brundage MD, Mackillop WJ. Perception of quantitative information for treatment decisions. Med Decis Making. 2000;20:228–38.
13. McAlister F, Straus S, Guyatt G, Haynes RB. Users' guide to the medical literature XX. Integrating research evidence with the care of the individual patient. J Am Med Assoc. 2000;283:2829–36.
14. Rothman A, Salovey P, Antone C, Keough K, Martin CD. The influence of message framing on intentions to perform health behaviors. J Exp Social Psychol. 1993;29:408–33.
15. Bobbio M, Demichelis B, Giustetto G. Completeness of reporting trial results: effect on physicians' willingness to prescribe. Lancet. 1994;343:1209–11.
16. Cranney M, Walley T. Same information, different decisions: the influence of evidence on the management of hypertension in the elderly. Br J Gen Pract. 1996;46:661–3.
17. Forrow L, Taylor WC, Arnold RM. Absolutely relative: how research results are summarized can affect treatment decisions. Am J Med. 1992;92:121–4.
18. Bucher HC, Weinbacher M, Gyr K. Influence of method of reporting study results on decision of physicians to prescribe drugs to lower cholesterol concentration. BMJ. 1994;309:761–4.
19. Naylor CD, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? Ann Intern Med. 1992;117:916–21.
20. Nikolajevic-Sarunac J, Henry DA, O'Connell DL, Robertson J. Effects of information framing on the intentions of family physicians to prescribe long-term hormone replacement therapy. J Gen Intern Med. 1999;14:591–8.
21. National Center for Education Statistics (US). Adult literacy in America: a first look at the findings of the national adult literacy survey. Washington, DC: US Department of Education; 1993.