

Different Ways to Describe the Benefits of Risk-Reducing Treatments

A Randomized Trial

Peder A. Halvorsen, MD; Randi Selmer, PhD; and Ivar Sønnebø Kristiansen, MD, PhD, MPH

Background: How physicians communicate the risks and benefits of medical care may influence patients' choices. Ways to communicate the benefits of risk-reducing drug therapies include the number needed to treat (NNT) to prevent adverse events, such as heart attacks or hip fractures, and gains in disease-free life expectancy or postponement of adverse events. Previous studies suggest that the magnitude of the NNT does not affect a layperson's decision about risk-reducing interventions, but postponement of an adverse event does affect such decisions.

Objective: To examine laypersons' responses to scenarios that describe benefits as postponing an adverse event or the equivalent NNT.

Design: Cross-sectional survey with random allocation to different scenarios.

Setting: General community.

Participants: Respondents to a population-based health study.

Intervention: The survey presented scenarios regarding a hypothetical drug therapy to reduce the risk for heart attacks (1754 respondents) or hip fractures (1000 respondents). The data sources for both scenarios were clinical trials. Respondents were randomly assigned to a scenario with 1 of 3 outcomes after 5 years of treatment. For the drug to prevent heart attacks, the outcomes were postponement by 2 months for all patients, postponement by 8 months for 1 of 4 patients, or an NNT of 13 patients to prevent 1 heart attack. For the drug to prevent hip fractures, the outcomes

were postponement by 16 days for all patients, postponement by 16 months for 3 of 100 patients, or an NNT of 57 patients to prevent 1 fracture.

Measurements: Consent to receive the intervention and perceived ease of understanding the treatment effect.

Results: The overall rate of response to the survey was 81%. In the heart attack scenarios, 93% of respondents who were presented with the NNT outcome consented to drug therapy, 82% who were presented with the outcome of large postponement for some patients consented to therapy, and 69% who were presented with the outcome of short postponement for all patients consented to therapy (chi-square, 89.6; $P < 0.001$). Corresponding consent rates for the hip fracture scenarios were 74%, 56%, and 34%, respectively (chi-square, 91.5, $P < 0.001$). Respondents who said that they understood the treatment effect were more likely to consent to therapy.

Limitation: Decisions were based on hypothetical scenarios, not real clinical encounters.

Conclusions: Treatment effects expressed in terms of NNT yielded higher consent rates than did those expressed as equivalent postponements. This result suggests that the description of the anticipated outcome may influence the patient's willingness to accept a recommended intervention.

Ann Intern Med. 2007;146:848-856.

www.annals.org

For author affiliations, see end of text.

Considerable resources are devoted to drug therapies that are aimed at modifying risk factors, such as hypertension, elevated cholesterol levels (1), and osteoporosis. For individual patients, the choice to begin preventive drug therapy should be consistent with their values and preferences. Thus, to engage meaningfully in shared decision making and to provide truly informed consent, patients need to have a clear understanding of the benefits and harms of a treatment. Strong and consistent evidence shows that stated preferences for medical interventions may

depend on how the treatment effects are described. For example, the likelihood of choosing a therapy may depend on whether its benefits are presented as absolute risk reductions or relative risk reductions (2) or as losses versus gains (3–5). These effects suggest the potential for influencing the patient's response by describing treatment effects in a certain way. We explore laypersons' responses to different ways of explaining possible outcomes of an intervention.

When informing decision makers about the benefit of risk-reducing drug therapies, several authors have advocated using the number needed to treat (NNT) to avoid 1 outcome (6–10), which is defined as the reciprocal of the absolute risk reduction. The NNT is the average number of patients in an intervention group who must be treated for a specific period to observe 1 fewer adverse outcome by the end of this period compared with those in a control group. Several authors believe that NNT provides an easily understood way to describe the effort needed to prevent adverse outcomes (9–11). However, for drug therapies aimed at disease processes that develop slowly, such as atherosclerosis and osteoporosis, the term *prevention* may be misleading. Rather than completely preventing adverse outcomes in a small fraction of patients, an intervention

See also:

Print

Editors' Notes	849
Glossary	855
Editorial comment.	891
Summary for Patients.	1-50

Web-Only

Appendix Tables
Conversion of figures and tables into slides

may postpone the event for many treated patients. Describing the outcome of treatment in terms of postponing an event may be a good alternative to using NNT for helping patients to understand the potential consequences of a decision.

We hypothesized that when laypersons consider preventive drug therapies, they will find the concept of time—and, hence, postponements—more useful than the concept of NNT. Specifically, we tested the hypotheses that laypersons perceive information about postponements as being easier to understand than the concept of NNT and that the rates at which a person consents to hypothetical drug therapies may depend on the measure that is used to describe the drugs' effects.

METHODS

Participants

In 2002, as part of regional health surveys in Norway (12, 13), the Norwegian Institute of Public Health invited all inhabitants born in 1925 to 1947, 1957, 1962, and 1972 and all persons born in 1948 to 1968 who had been invited to former screenings in Finnmark County, Norway, to participate in the Troms and Finnmark (TRO-FINN) health study (13). For each participant, blood pressure and body mass index were measured and a blood sample was drawn to measure lipids and glucose levels. The participants completed 2 comprehensive questionnaires that included sociodemographic data; health-related information; and habits regarding exercise, food and alcohol consumption, and smoking. Two weeks after screening, the participants received a letter with their results. Participants who were at high risk for cardiovascular disease were advised to contact their general practitioner for follow-up.

For our study, we surveyed a sample of participants from the TROFINN study (who lived in 10 municipalities along the coast of Finnmark) about their preferences for risk-reducing drug therapies. **Figure 1** shows the formation of our study sample. Of the 11 284 persons invited to screening, 6854 (61%) participated, of whom 6445 were eligible for our study. Eligibility was based on the person's written consent to allow researchers to use data from the initial screening and his or her willingness to be approached about future surveys. We excluded persons who died or emigrated between the initial screening and the date of our survey and those with a missing address. For other study purposes (14), we ranked the participants according to their cardiovascular risk to identify high-risk ($n = 754$) and low-risk ($n = 1000$) persons. We surveyed these persons about whether they would use a hypothetical drug aimed at reducing the risk for a heart attack. We also surveyed a random sample of the remaining persons ($n = 1000$) about whether they would use a hypothetical drug to reduce the risk for hip fracture. To maximize the response rate, we mailed 1 reminder letter and a copy of the questionnaire to nonresponders. We planned to enroll ap-

Context

In previous research, different ways of describing the outcomes of an intervention led to different health care decisions.

Contribution

Healthy people were randomly assigned to receive equal but different descriptions of the outcome of a hypothetical intervention to prevent myocardial infarction (MI). Responders were more likely to consent to treatment when the outcome was described as the number needed to treat to prevent 1 MI. They were less likely to consent when the intervention was described as not preventing but delaying an MI by 2 months for all persons or by 8 months for 25% of persons.

Caution

The scenarios were presented in a survey and were hypothetical.

Implication

Quantitatively equal but differently worded outcomes elicit different health care decisions.

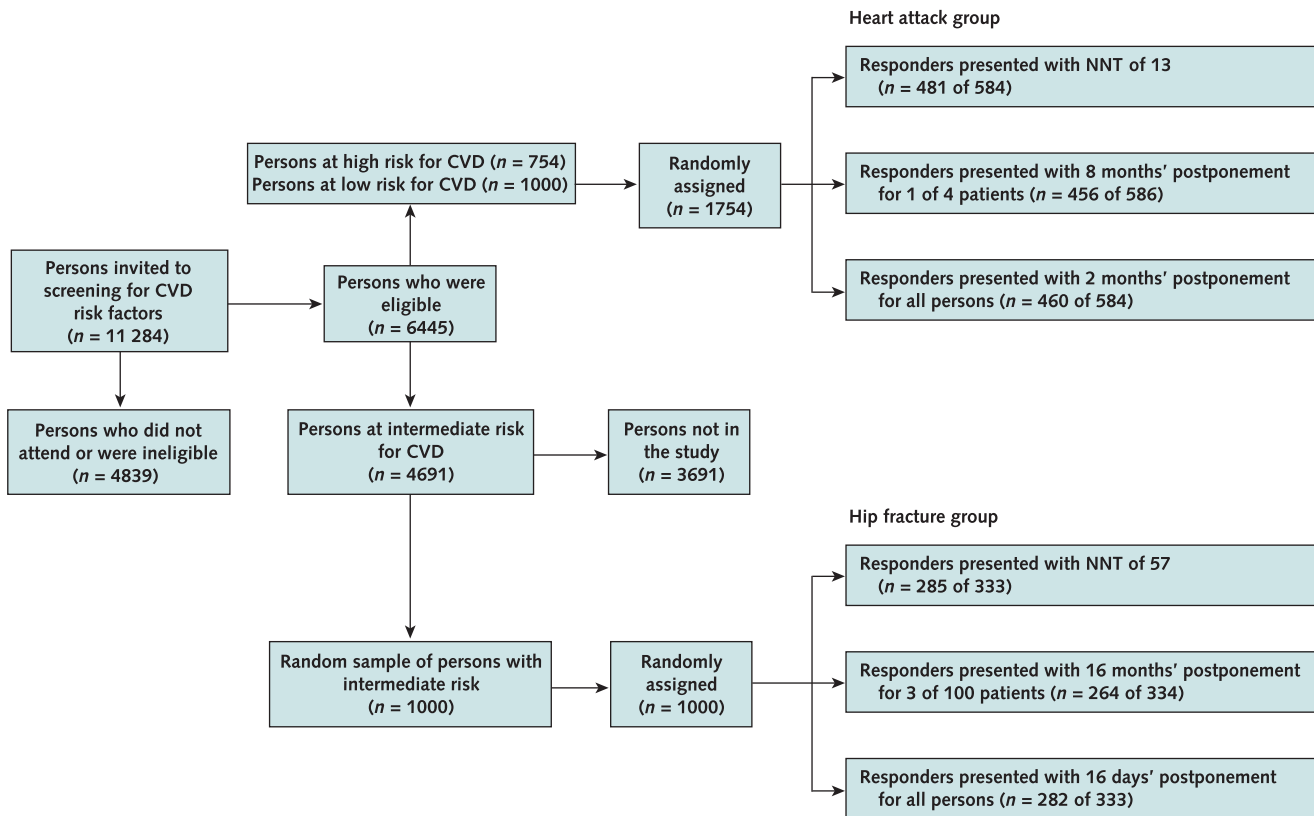
—The Editors

proximately 1000 persons in each group, but the number of high-risk persons was lower than expected.

Procedures

We presented a scenario that described hypothetical drug therapy to prevent a heart attack to persons in 1 group and a scenario that described such therapy to prevent hip fracture to persons in the other group (**Figure 2**). Using data from the Scandinavian Simvastatin Survival Study (15, 16), we calculated the NNT to prevent 1 heart attack after 5 years of therapy (NNT, 13) and the number of participants for whom treatment would postpone a heart attack and the length of the disease-free interval (a 2-month average postponement for all patients and an 8-month postponement for 1 of 4 patients, with no benefit for 3 of 4 patients) (**Figure 2**; **Appendix Tables 1 and 2**, available at www.annals.org). We used a computerized random-sample function (SPSS, Chicago, Illinois) to randomly assign respondents in the high-risk and low-risk groups to 1 of the 3 scenarios (**Figure 1**). We also calculated the benefit of hip fracture prevention after 5 years of therapy using data from the Fracture Intervention Trial (17). The NNT to prevent 1 hip fracture was 57. Alternatively, as a result of treatment, all patients would have a hip fracture 16 days later than they would have without treatment or only 3 of 100 patients would have a hip fracture 16 months later than they would have without treatment, whereas the remaining patients would not benefit (**Figure 2**; **Appendix Tables 1 and 2**, available at www.annals.org). Again, allocation to the 3 scenarios was random. One survey question asked respondents whether they found it very easy, somewhat

Figure 1. Study flow diagram.



A total of 2754 attendees to a population-based health study were randomly assigned to hypothetical scenarios that presented the benefits of preventive drug therapies in terms of number needed to treat (NNT) or postponement of adverse events. Eligibility criteria were as follows: attended screening, consented to additional studies, were alive or did not emigrate between the time of the screening and the survey, and had a known address. Strategic allocation to the study groups by risk for cardiovascular disease (CVD) was done for other study purposes, and for similar reasons, the low-risk sample was selected so that the proportion of women was the same as that in the high-risk sample (14). We expected to enroll approximately 1000 persons in each risk group, but the number of high-risk persons was lower than expected.

easy, somewhat difficult, or very difficult to understand the treatment effect. Another question asked whether the respondent would consent to the hypothetical drug therapy because of the benefits described in the scenario. Possible response categories were “certainly,” “probably,” “probably not,” or “certainly not.” We linked each respondent’s answers to his or her responses to the health survey given by the Norwegian Institute of Public Health.

Outcome Measures

We tested the hypotheses that more respondents would report difficulties in understanding the NNT effect format than the 2 postponement formats and that the proportion of respondents who consented to therapy would differ among the 3 scenarios. Therefore, the primary outcome measures were the rates of consent to therapy and the difficulty in understanding the effect format. Although the survey provided the respondent with several graded responses, we dichotomized these variables when we analyzed the survey data. Therefore, we defined difficulty as a response of “very difficult” or “quite difficult.” We defined consent as a response of “certainly” or “probably” to the

question regarding willingness to consent to the therapy. Analysis of these variables as 4-point responses instead of as dichotomous responses yielded similar results, but for ease of understanding, we present the results with the responses grouped as described. To test whether difficulty understanding the outcome measure had an effect on consent rates, we analyzed the perceived difficulty of understanding as a possible predictor of consent to therapy. We used age, sex, level of education, self-reported health state, smoking habits, and psychiatric symptoms as secondary independent variables in both study groups. In the heart attack group, history of cardiovascular diseases, use of lipid-lowering agents or antihypertensive medications, and premature coronary heart disease among close relatives were additional secondary variables. We used a history of fractures as a secondary variable in the hip fracture group.

Statistical Analysis

We evaluated differences between proportions by using chi-square tests. We used log-Poisson regression with robust SEs to explore possible associations (expressed as relative risks) between the dependent variables (primary

outcome measures) and independent variables other than effect format. We tested for first-order interactions between effect format and other independent variables by adding product terms to the regression models. We also tested for interaction between level of education and perceived difficulty of understanding in predicting consent to therapy (dependent variable). We did not perform formal power calculations when we designed the study. We used SPSS, version 10.0 (SPSS), and Stata, version 9.2 (Stata Corp., College Station, Texas). The Norwegian Data Inspectorate approved the TROFINN study, and the regional committee for medical research ethics evaluated the study. We conducted the study in accordance with the Declaration of Helsinki.

Role of the Funding Source

The funding body had no involvement in the design of the study, data collection and interpretation, or decision to submit the manuscript for publication.

RESULTS

Study Sample

In the heart attack group, 1397 of 1754 (80%) participants responded to the questionnaire, whereas 831 of 1000 (83%) participants in the hip fracture group responded. In the respective study groups, the mean age of the participants was 58 years (SD, 11) and 60 years (SD, 10) and the proportion of women was 34% and 60%. Nonresponders did not differ substantially from responders with respect to age and sex. In the heart attack group, responders were slightly better educated than were nonresponders (10.4 years [95% CI, 10.2 to 10.6 years] vs. 9.9 years [CI, 9.5 to 10.3 years]). In the hip fracture group, more responders than nonresponders (62% [CI, 59% to 65%] vs. 53% [CI, 45% to 61%]) reported good health. In

Figure 2. The questionnaire for the heart attack group.

Imagine that your doctor informs you that your risk for heart attack is elevated. The doctor suggests that you take a drug to prevent a heart attack. The drug must be taken daily, and you will have to visit your doctor twice a year for a check-up. Side effects are neither common nor dangerous. Drug costs are refunded according to the rules of the National Health Insurance.

Q1a. The doctor informs you that for every heart attack that is prevented, 13 patients have to take the drug for 5 years.

or

Q1b. The drug therapy may not completely prevent heart attacks. Rather, it postpones heart attacks for a while. The doctor informs you that all patients who take the drug therapy for 5 years will live about 2 months longer before they get a heart attack.

or

Q1c. The drug therapy may not completely prevent heart attacks. Rather, it postpones heart attacks for a while. The doctor informs you that 1 of 4 patients who take the drug for 5 years will live about 8 months longer before they get a heart attack, whereas the others will have no benefit from the drug therapy.

Would you choose to take this drug?

Certainly Probably Probably not Certainly not

Q2. Did you find it easy or difficult to understand the magnitude of the effect of this drug therapy?

Very easy Somewhat easy Somewhat difficult Very difficult

In each questionnaire, only 1 of the 3 versions of item Q1 (a, b, or c) was used. The respondents were randomly allocated to 1 version of the questionnaire only. We used similar scenarios with different numbers for the hip fracture questionnaire.

Table 1. Characteristics of Respondents*

Variable	Scenario		
	NNT to Prevent 1 Outcome (95% CI)	Long Postponement of an Outcome for Some Patients (95% CI)	Short Postponement of an Outcome for All Patients (95% CI)
Heart attack group			
Mean age, y	58 (57–59)	58 (57–59)	57 (56–58)
Women, %	32 (28–36)	37 (32–41)	34 (30–38)
Education >12 y, %	22 (18–25)	23 (19–27)	24 (20–28)
Good self-reported health, %	61 (56–65)	62 (57–66)	63 (58–67)
History of cardiovascular diseases, %	21 (18–25)	20 (16–24)	21 (17–24)
Hip fracture group			
Mean age, y	60 (59–61)	60 (59–61)	60 (59–61)
Women, %	56 (50–62)	63 (57–69)	62 (57–68)
Education >12 y, %	19 (14–24)	20 (15–25)	20 (15–24)
Good self-reported health, %	61 (56–67)	60 (54–66)	65 (59–70)
History of fractures, %	15 (11–19)	11 (7–15)	13 (9–17)

* For the heart attack group, there were 481, 456, and 460 respondents in the NNT scenario, long postponement of an outcome scenario, and short postponement of an outcome scenario, respectively. For the hip fracture group, these numbers were 285, 264, and 282, respectively. Missing responses were 4% or less for all categories. For some categories, the number of respondents may therefore be slightly lower than indicated. NNT = number needed to treat.

Table 2. Rates of Consent to Therapy and Difficulties with Understanding the Treatment Effect*

Variable	NNT to Prevent 1 Outcome	Long Postponement of Outcome for Some Patients	Short Postponement of Outcome for All Patients
Persons in the heart attack group, n (%)			
Would you choose to take this drug?			
Certainly	273 (47)	190 (32)	153 (26)
Probably	172 (29)	179 (31)	162 (28)
Probably not	24 (4)	66 (11)	94 (16)
Certainly not	10 (2)	17 (3)	49 (8)
Missing responses or nonresponders	105 (18)	134 (23)	126 (22)
Was it difficult or easy to understand the magnitude of the effect of this drug therapy?			
Very easy	96 (16)	80 (14)	85 (15)
Somewhat easy	226 (39)	219 (37)	189 (32)
Somewhat difficult	133 (23)	131 (22)	145 (25)
Very difficult	25 (4)	22 (4)	37 (6)
Missing responses or nonresponders	104 (18)	134 (23)	128 (22)
Persons in the hip fracture group, n (%)			
Would you choose to take this drug?			
Certainly	78 (23)	50 (15)	27 (8)
Probably	132 (40)	98 (29)	69 (21)
Probably not	48 (14)	77 (23)	89 (27)
Certainly not	25 (8)	39 (12)	96 (29)
Missing responses or nonresponders	50 (15)	70 (21)	52 (16)
Was it difficult or easy to understand the magnitude of the effect of this drug therapy?			
Very easy	36 (11)	50 (15)	58 (17)
Somewhat easy	81 (24)	81 (24)	76 (23)
Somewhat difficult	114 (34)	96 (29)	88 (26)
Very difficult	51 (15)	37 (11)	59 (18)
Missing responses or nonresponders	51 (15)	70 (21)	52 (16)

* For the heart attack group, there were 584, 586, and 584 participants in the NNT scenario, long postponement of an outcome scenario, and short postponement of an outcome scenario, respectively. For the hip fracture group, these numbers were 333, 334, and 333, respectively. NNT = number needed to treat.

both groups, participants in the 3 scenarios were fairly balanced in age, sex, education, self-assessed overall health, and history of cardiovascular diseases or fractures (Table 1).

Responses to the Scenarios

Table 2 shows the numbers of responses for each study group. In both groups, the proportion of respondents who consented to therapy was highest when the treatment effects were presented in terms of NNT, intermediate when treatment effects were presented as a long postponement of the outcome for a fraction of the patients, and lowest when treatment effects were presented as a short postponement of the outcome for all patients (Table 3). The differences were statistically significant in the heart attack group (93% [CI, 91% to 95%] vs. 82% [CI, 78% to 85%] vs. 69% [CI, 65% to 73%]; chi-square, 90; $P < 0.001$) and in the hip fracture group (74% [CI, 69% to 78%] vs. 56% [CI, 50% to 62%] vs. 34% [CI, 29% to 40%]; chi-square, 92; $P < 0.001$). In the regression analysis, greater perceived ease of understanding the effect measure and less education were additional independent predictors of consent to therapy (Table 3). Many persons in both study groups had difficulty understanding the treatment effect (Table 4), but differences among the scenarios were not statistically significant. Regression analysis of the hip fracture data—but not the heart attack data—indi-

cated that difficulties were greater among older persons and those with less education (Table 4).

DISCUSSION

In this population-based survey, laypersons were more inclined to accept therapy to reduce the risk for heart attack or hip fracture when the benefit was presented as the NNT to prevent 1 adverse outcome than when presented as postponements of the outcome. The benefits described in all 3 scenarios were equivalent because we used the same clinical study to calculate them. Many respondents reported difficulty understanding the description of treatment benefit regardless of how we presented it, and such persons were less likely to consent to therapy. These findings are intriguing when placed in the context of informed consent, patient-directed choices, and shared decision making. Because assisting patients in decision making is a core element of the physician’s work, knowing that decisions may be influenced by the words used to describe benefits, and perhaps harms, is important for clinical practice. The main body of empirical knowledge, however, stems from the field of experimental cognitive psychology.

Seminal works in cognitive psychology (3) and medical decision making (4) have emphasized that a person’s

Table 3. Respondents' Consents to a Drug Therapy Aimed at Preventing Heart Attacks or Hip Fractures*

Variable	Heart Attack Group		Hip Fracture Group	
	Respondents Consenting to the Drug Therapy, n/n (%)	Relative Risk (95% CI)	Respondents Consenting to the Drug Therapy, n/n (%)	Relative Risk (95% CI)
Effect measure				
Postponement: same benefit for all persons (reference)	315/458 (69)	1.00	96/281 (34)	1.00
Postponement: greater benefit for some persons	369/452 (82)	1.17 (1.08–1.27)	148/264 (56)	1.60 (1.31–1.96)
NNT	445/479 (93)	1.34 (1.25–1.44)	210/283 (74)	2.16 (1.80–2.59)
Perception of effect measure				
Easy to understand	792/894 (89)	1.28 (1.20–1.37)	235/382 (62)	1.36 (1.21–1.54)
Difficult to understand (reference)	335/490 (68)	1.00	216/442 (49)	1.00
Age				
		1.03 (1.00–1.06) [†]		1.04 (0.96–1.11) [†]
<50 y (reference)	192/266 (72)		58/117 (50)	
50–59 y	398/485 (82)		150/281 (53)	
60–69 y	351/418 (84)		143/265 (54)	
≥70 y	188/220 (86)		103/165 (62)	
Sex				
Women (reference)	378/476 (79)	1.00	269/499 (54)	1.00
Men	751/913 (82)	1.00 (0.94–1.06)	185/329 (56)	1.00 (0.88–1.14)
Length of education				
		0.96 (0.93–1.00) [†]		0.85 (0.77–0.93) [†]
0–9 y (reference)	526/620 (85)		250/406 (62)	
10–12 y	347/428 (81)		123/241 (51)	
≥12 y	230/310 (74)		69/159 (43)	
Self-reported health condition				
Not good or bad	434/528 (82)	0.96 (0.91–1.02)	186/311 (60)	1.08 (0.95–1.23)
Good or excellent (reference)	687/853 (81)	1.00	268/515 (52)	1.00
Psychiatric symptoms				
Yes (reference)	611/761 (80)	1.00	288/508 (57)	1.00
No	452/551 (82)	1.02 (0.96–1.07)	145/282 (51)	0.95 (0.83–1.10)
Smoking habits				
Present smoker	318/375 (85)	1.09 (1.02–1.17)	162/281 (58)	1.15 (0.97–1.36)
Former smoker	469/566 (83)	1.06 (0.99–1.13)	173/302 (57)	1.08 (0.92–1.28)
Never smoked (reference)	336/441 (76)	1.00	116/242 (48)	1.00
Previous fractures				
Yes	–	–	53/104 (51)	0.86 (0.72–1.04)
No (reference)	–	–	386/701 (55)	1.00
Cardiovascular disease				
Yes	240/282 (85)	0.99 (0.92–1.07)	–	–
No (reference)	870/1087 (80)	1.00	–	–
Family history of premature cardiovascular disease				
Yes	367/426 (86)	1.05 (0.99–1.10)	–	–
No (reference)	762/963 (80)	1.00	–	–
Taking lipid-lowering drugs				
Yes	284/330 (86)	1.01 (0.94–1.08)	–	–
No (reference)	824/1035 (80)	1.00	–	–
Taking antihypertensive drugs				
Yes	349/406 (86)	1.05 (0.98–1.11)	–	–
No (reference)	774/976 (79)	1.00	–	–

* Multivariate Poisson regression analysis was used to estimate relative risks. Respondents were considered to have consented if they indicated that they would “certainly” or “probably” accept the therapy. Otherwise, consent was considered absent. NNT = number needed to treat.
[†] Trend analysis across the subgroups.

decisions may depend on how the outcomes of interventions are described. The 2 descriptions in each of the following pairs may evoke different choices: *losses versus gains*

(see **Glossary**) (3, 4), gains in *life expectancy* (see **Glossary**) versus gains in *cumulative probability* (see **Glossary**) of survival (4), and *certain outcomes* (see **Glossary**) versus *uncer-*

tain outcomes (see **Glossary**) (18). For example, in the context of lung cancer treatment, McNeil and coworkers (4) found that surgery was more attractive to respondents when the benefits were described in terms of life expectancy than when they were expressed as cumulative probabilities of survival. Collectively, these differences are known as “framing effects”. Several studies have shown that persons choose differently when the outcome is framed as an NNT or absolute risk reduction rather than as a relative risk reduction (19–22). In a recent study (23), laypersons considering a hypothetical drug therapy for osteoporosis discriminated between levels of effectiveness that were presented as degrees of postponement of hip fracture but not when presented in terms of different NNTs.

Why, then, did we observe higher consent rates when we expressed treatment effects as NNTs rather than as postponements of an outcome? First, the findings may reflect preferences for a large but uncertain benefit (an outcome avoided for only a few patients) over a smaller benefit enjoyed by all patients. Respondents may have perceived 2 of the effect formats as gambles (24): 1 with a “big prize”

of completely avoiding the adverse outcome (the NNT format) and the other with a substantial, but not indefinite, postponement of the adverse outcome for a few patients. The hypothesis that persons perceive the NNT format as a gamble has empirical support from surveys of laypersons (23–25) and physicians (26). Tversky and Kahneman (3) observed that persons tend to be risk-averse (preferring a certain, intermediate outcome) when outcomes are described as gains, but risk-seeking (preferring a gamble to get a better outcome) when they perceive the outcomes as worsening their status (3). Eraker and Sox (18) replicated this finding with medical scenarios. By analogy, perhaps our respondents are risk-seeking because the risky effect formats (NNT and long postponement of an outcome for some patients) evoked the highest consent rates. We framed our scenarios as an imminent loss of health (**Figure 2**)—the prospect of a heart attack or a hip fracture—which may explain why treatment effects framed as gambles were attractive to our respondents.

A second explanation is that different consent rates with different outcome descriptions may reflect what cog-

Table 4. Respondents Reporting Difficulties with Understanding the Benefits of a Preventive Drug Therapy*

Variable	Heart Attack Group		Hip Fracture Group	
	Respondents Reporting Difficulties, n/n (%)	Relative Risk (95% CI)	Respondents Reporting Difficulties, n/n (%)	Relative Risk (95% CI)
Effect measure				
Postponement: same benefit for all persons (reference)	182/456 (40)	1.00	147/281 (52)	1.00
Postponement: greater benefit for some persons	153/452 (34)	0.85 (0.71–1.01)	133/264 (50)	0.98 (0.83–1.15)
NNT	158/480 (33)	0.84 (0.70–1.00)	165/282 (59)	1.12 (0.96–1.31)
Age				
<50 y (reference)	94/266 (35)	1.00	45/117 (38)	1.00
50–59 y	171/484 (35)	0.99 (0.79–1.24)	150/282 (53)	1.35 (1.05–1.75)
60–69 y	140/420 (33)	0.91 (0.71–1.16)	150/265 (57)	1.44 (1.10–1.87)
≥70 y	88/218 (40)	1.06 (0.81–1.40)	100/163 (61)	1.57 (1.18–2.08)
Sex				
Women (reference)	175/474 (37)	1.00	267/500 (53)	1.00
Men	318/914 (35)	0.97 (0.83–1.14)	178/327 (54)	0.98 (0.86–1.13)
Length of education				
0–9 y (reference)	225/618 (36)	1.00	233/405 (58)	1.00
10–12 y	137/428 (32)	0.86 (0.71–1.03)	132/241 (55)	1.03 (0.89–1.20)
≥12 y	114/310 (37)	0.99 (0.80–1.23)	68/159 (43)	0.80 (0.65–1.00)†
Self-reported health condition				
Not good or bad	190/528 (36)	0.98 (0.83–1.17)	162/312 (52)	0.88 (0.76–1.02)
Good or excellent (reference)	300/582 (35)	1.00	282/513 (55)	1.00
Psychiatric symptoms				
Yes	182/551 (33)	1.15 (0.97–1.35)	273/507 (54)	1.05 (0.91–1.22)
No (reference)	285/761 (37)	1.00	148/282 (52)	1.00
Smoking habits				
Present smoker	124/375 (33)	0.92 (0.75–1.13)	151/280 (54)	1.02 (0.86–1.21)
Former smoker	202/566 (36)	1.00 (0.84–1.20)	159/301 (53)	0.94 (0.80–1.11)
Never smoked (reference)	164/439 (37)	1.00	133/243 (55)	1.00

* Multivariate Poisson regression analysis was used to estimate relative risks. NNT = number needed to treat.

† P = 0.048.

nitive psychology teaches us about the properties of *heuristics* (see **Glossary**). When using these shortcuts in making decisions, persons may put the most weight on cues that are easily recalled (*availability heuristics* [see **Glossary**] [27]), easily recognized (recognition heuristics [28]), or easily evaluated (the *evaluability hypothesis* [see **Glossary**] [29]). Our scenarios portrayed some cues that persons should easily recognize: a serious disease (heart attack or hip fracture), the word *prevention*, and the word *postponement*. Our scenarios also portrayed some outcomes as numbers: the NNT to prevent 1 event and the length of a postponement. One reason for using a written scenario is to standardize the stimulus to the respondent. Numbers can be a source of unwanted variability if they are evaluated differently. In fact, persons vary in their numeracy and may have difficulty evaluating a number, such as an NNT, without a reference point for comparison. Under such circumstances, the evaluability hypothesis (29) predicts that NNTs would have little impact on decisions, a finding that has considerable empirical support (23–25, 30). However, one would expect laypersons to be more familiar with the concept of time and thus to be better able to evaluate the extent of postponement of an adverse outcome. Perhaps some respondents simplified the scenarios to “complete prevention of heart attack” (in the NNT scenarios), “substantial postponement of heart attack,” and “small postponement of heart attack.” This heuristic simplifies the decision-making task by circumventing the need to interpret numbers. The rank ordering of these 3 simplified options according to their face value is the same as the observed rates of consent with the 3 scenarios (**Table 3**).

A third explanation is that the NNT and postponement scenarios have different face validity in representing the true treatment effect. The NNT is like a lottery—a few people win a “big prize” and the rest receive nothing. However, the NNT format leaves much unsaid. One cannot directly infer the proportion of patients who benefit from the intervention (24) or that the “lucky ones” will completely avoid an adverse outcome. Of the 2 postponement formats, the “greater postponement for a proportion of patients” format is similar to the outcomes of the studies on which we based the scenarios (16, 17) (**Appendix Tables 1 and 2**, available at www.annals.org). In these studies, most participants in the control groups did not experience the adverse event. Thus, only a minority of participants could benefit during the study period. In our hypothetical scenarios, we did not state the proportion of persons who would have an adverse event without the drug therapy. In the end, however, neither hip fractures nor heart attacks are inevitable. The NNT format reflects this uncertainty, albeit indirectly, whereas the short postponement for all patients do not. Thus, it is conceivable that the consent rates were higher in the NNT scenarios because they seemed more plausible to the participants.

This study has important strengths—proper randomization, a large community-based survey sample, and high

Glossary

- Availability heuristics*: The process by which people judge an event as more likely if it is easily recalled. Judgment based on what comes easily to mind.
- Certain outcomes*: Outcomes that are inevitable (that is, $P = 1.0$) (for example, death, at least in the long run).
- Cumulative probability*: The probability that an event has occurred at a specific point in time (for example, 10 years after the onset of an intervention). A gain in cumulative probability may be presented in several ways, such as relative risk reduction, absolute risk reduction, or number needed to treat.
- Evaluability hypothesis*: This hypothesis states that when making complex choices, people tend to base their decisions on cues or factors that are easy to evaluate and put little or no weight on factors that are hard to evaluate.
- Heuristics*: Mental short cuts or “rules of thumb” that people use to simplify complex decision-making processes.
- Life expectancy*: The average statistical (expected) remaining lifetime for a group of people with similar characteristics (for example, 60-year-old men with localized prostate cancer). Gains in life expectancy is the extent to which a medical intervention, on average, extends (disease-free) life.
- Losses versus gains*: Logically equivalent ways of presenting the outcome of an intervention in terms of losses (for example, 1 of 100 people will die) or gains (99 of 100 people will survive).
- Uncertain outcomes*: Outcomes that may or may not occur (that is, $P < 1.0$) (for example, heart attack, stroke, or cancer).

response rates—and several limitations. First, although the participants were selected from respondents to a general health survey, only 61% of invited persons attended the screening, which means that our respondents may not be entirely representative of the general population. Second, the survey questionnaire did not probe deeply into respondents’ understanding of the scenarios. Qualitative studies or more direct assessments of understanding (31) might teach us more about why persons report difficulty in understanding measures of the effect of preventive drug therapies. Third, whether responses to hypothetical scenarios reflect real-life decisions is a long-standing concern about studies such as ours (32, 33). We cannot rule out the possibility that the respondents’ reactions to the outcomes described in the scenarios have little resemblance to real-life decisions. In particular, our scenarios did not include measures of the uncertainty in the effect estimates. Henderson and Keiding (34) have proposed methods for communicating uncertainty in effect estimates of survival, but to our knowledge, no studies have empirically tested these strategies. Fourth, our scenarios did not specify potential harms or represent the tradeoffs between benefits and harms. Finally, we did not tailor the scenarios to the individual respondents’ level of risk. Therefore, the respondents were free to use assessment of their own risk, which may have introduced unwanted variability in their responses. Including measures of baseline risks, uncertainty, or harms to our scenarios would have added realism and complexity to the decision tasks. Respondents might have reacted by relying more on heuristics that focused on the salient attributes of the scenarios, such as the seriousness of the disease to be prevented (30).

Notwithstanding these limitations, we conclude that the NNT format and the 2 postponement formats, as dif-

difficult as they may be to understand, evoke very different choices when laypersons respond to hypothetical scenarios about risk-reducing drug therapies. Perhaps the same is true in real life.

From the University of Southern Denmark, Odense, Denmark; University of Tromsø, Tromsø, Norway; and Norwegian Institute of Public Health and University of Oslo, Oslo, Norway.

Acknowledgments: The authors thank Arthur Elstein, Dorte Gyrd-Hansen, Jørgen Nexøe, and Henrik Støvring for comments on the paper. The authors also thank the residents of Finnmark, Norway, for their willingness to contribute to the study and the Norwegian Institute of Public Health for conducting the practical part of the data collection.

Grant Support: By governmental funds held by the University of Tromsø, Tromsø, Norway.

Potential Financial Conflicts of Interest: None disclosed.

Requests for Single Reprints: Peder A. Halvorsen, MD, Nordlys Legesenter, N-9509 Alta, Norway; e-mail, phalvor@online.no.

Current author addresses and author contributions are available at www.annals.org.

References

1. Getz L, Sigurdsson JA, Hetlevik I, Kirkengen AL, Romundstad S, Holmen J. Estimating the high risk group for cardiovascular disease in the Norwegian HUNT 2 population according to the 2003 European guidelines: modelling study. *BMJ*. 2005;331:551. [PMID: 16103030]
2. Edwards A, Elwyn G, Covey J, Matthews E, Pill R. Presenting risk information—a review of the effects of “framing” and other manipulations on patient outcomes. *J Health Commun*. 2001;6:61-82. [PMID: 11317424]
3. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science*. 1981;211:453-8. [PMID: 7455683]
4. McNeil BJ, Pauker SG, Sox HC Jr, Tversky A. On the elicitation of preferences for alternative therapies. *N Engl J Med*. 1982;306:1259-62. [PMID: 7070445]
5. Ghosh AK, Ghosh K. Translating evidence-based information into effective risk communication: current challenges and opportunities. *J Lab Clin Med*. 2005;145:171-80. [PMID: 15962835]
6. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med*. 1988;318:1728-33. [PMID: 3374545]
7. Gigerenzer G, Edwards A. Simple tools for understanding risks: from innu-meracy to insight. *BMJ*. 2003;327:741-4. [PMID: 14512488]
8. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ*. 1995;310:452-4. [PMID: 7873954]
9. Riegelman R, Schroth WS. Adjusting the number needed to treat: incorporating adjustments for the utility and timing of benefits and harms. *Med Decis Making*. 1993;13:247-52. [PMID: 8412555]
10. Rajkumar SV, Sampathkumar P, Gustafson AB. Number needed to treat is a simple measure of treatment efficacy for clinicians. *J Gen Intern Med*. 1996;11:357-9. [PMID: 8803742]
11. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. Therapy. In: Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB, eds. *Evidence-based Medicine: How to Practice and Teach EBM*. 2nd ed. London: Churchill Livingstone; 2000:105-53.
12. Bjartveit K, Foss OP, Gjervig T. The cardiovascular-disease study in Norwegian counties. Results from first screening. *Acta Medica Scand*. 1983;675:1-184. [PMID: 1871734]
13. Sogaard AJ, Gustad TK, Bjertness E, Tell GS, Schei B, Emaus N, Meyer HE, et al. Urban-rural differences in distal forearm fractures: cohort Norway. *Osteoporos Int*. 2007 Mar 1. [Epub ahead of print]. [PMID: 17333447]
14. Halvorsen PA, Kristiansen IS, Selmer R. Does information about cardiovascular risk status affect people's perception of their own life expectancy [Abstract]? In: Abstracts from the European Society for Medical Decision Making Meeting Rotterdam, the Netherlands, 7–8 June 2004. *Med Decis Making*. 2004;24:423.
15. van Hout BA, Simoons ML. Cost-effectiveness of HMG coenzyme reductase inhibitors; whom to treat? *Eur Heart J*. 2001;22:751-61. [PMID: 11350107]
16. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet*. 1994;344:1383-9. [PMID: 7968073]
17. Black DM, Thompson DE, Bauer DC, Ensrud K, Musliner T, Hochberg MC, et al. Fracture risk reduction with alendronate in women with osteoporosis: the Fracture Intervention Trial. FIT Research Group. *J Clin Endocrinol Metab*. 2000;85:4118-24. [PMID: 11095442]
18. Eraker SA, Sox HC Jr. Assessment of patients' preferences for therapeutic outcomes. *Med Decis Making*. 1981;1:29-39. [PMID: 7052402]
19. Nexøe J, Gyrd-Hansen D, Kragstrup J, Kristiansen IS, Nielsen JB. Danish GPs' perception of disease risk and benefit of prevention. *Fam Pract*. 2002;19:3-6. [PMID: 11818342]
20. Sheridan SL, Pignone MP, Lewis CL. A randomized comparison of patients' understanding of number needed to treat and other common risk reduction formats. *J Gen Intern Med*. 2003;18:884-92. [PMID: 14687273]
21. 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. [PMID: 7950558]
22. Misselbrook D, Armstrong D. Patients' responses to risk information about the benefits of treating hypertension. *Br J Gen Pract*. 2001;51:276-9. [PMID: 11458479]
23. Christensen PM, Brosen K, Brixen K, Andersen M, Kristiansen IS. A randomized trial of laypersons' perception of the benefit of osteoporosis therapy: number needed to treat versus postponement of hip fracture. *Clin Ther*. 2003;25:2575-85. [PMID: 14667958]
24. Kristiansen IS, Gyrd-Hansen D, Nexøe J, Nielsen JB. Number needed to treat: easily understood and intuitively meaningful? Theoretical considerations and a randomized trial. *J Clin Epidemiol*. 2002;55:888-92. [PMID: 12393076]
25. Nexøe J, Kristiansen IS, Gyrd-Hansen D, Nielsen JB. Influence of number needed to treat, costs and outcome on preferences for a preventive drug. *Fam Pract*. 2005;22:126-31. [PMID: 15640298]
26. Halvorsen PA, Kristiansen IS, Aasland OG, Førde OH. Medical doctors' perception of the “number needed to treat” (NNT). A survey of doctors' recommendations for two therapies with different NNT. *Scand J Prim Health Care*. 2003;21:162-6. [PMID: 14531508]
27. Tversky A, Kahneman D. Availability: heuristic for judging frequency and probability. *Cognitive Psychology*. 1973;5:207-32.
28. Gigerenzer G, Todd P, ABC Research Group. The recognition heuristic: how ignorance makes us smart. In: Goldstein DG, Gigerenzer, eds. *Simple Heuristics That Make Us Smart*. New York: Oxford Univ Pr; 1999:37-58.
29. Hsee CK, Loewenstein GF, Blount S, Bazerman MH. Preference reversals between joint and separate evaluations of options: a review and theoretical analysis. *Psychol Bull*. 1999;125:576-90.
30. Halvorsen PA, Kristiansen IS. Decisions on drug therapies by numbers needed to treat: a randomized trial. *Arch Intern Med*. 2005;165:1140-6. [PMID: 15911727]
31. Sheridan SL, Pignone M. Numeracy and the medical student's ability to interpret data. *Eff Clin Pract*. 2002;5:35-40. [PMID: 11874195]
32. Finch J. The vignette technique in survey research. *Sociology*. 1987;21:105-14.
33. Jones TV, Gerrity MS, Earp J. Written case simulations: do they predict physicians' behavior? *J Clin Epidemiol*. 1990;43:805-15. [PMID: 2200852]
34. Henderson R, Keiding N. Individual survival time prediction using statistical models. *J Med Ethics*. 2005;31:703-6. [PMID: 16319233]

Current Author Addresses: Dr. Halvorsen: Nordlys Legesenter, N-9509 Alta, Norway.

Dr. Selmer: Division of Epidemiology, Norwegian Institute of Public Health, PO Box 4404, Nydalen, N-0403 Oslo, Norway.

Dr. Kristiansen: Institute of Health Management and Health Economics, University of Oslo, PO Box 1089, Blindern, N-0317 Oslo, Norway.

Author Contributions: Conception and design: P.A. Halvorsen, R. Selmer, I.S. Kristiansen.

Analysis and interpretation of the data: P.A. Halvorsen, I.S. Kristiansen. Drafting of the article: P.A. Halvorsen, I.S. Kristiansen.

Critical revision of the article for important intellectual content: R. Selmer, I.S. Kristiansen.

Final approval of the article: P.A. Halvorsen, R. Selmer, I.S. Kristiansen. Statistical expertise: R. Selmer.

Appendix Table 1. Estimation of Corresponding Magnitudes of Number Needed to Treat and Postponement of Adverse Events after 5 Years of Drug Therapy in the Heart Attack Group*

Variable	Data	
	Placebo Group	Intervention Group
Data from the 4S study according to van Hout and Simoons (15)		
Participants, <i>n</i>	2223	2221
Events (stroke or heart attack), <i>n</i>	576	400
Mean event-free survival, <i>y</i>	4.32	4.48
Our estimations		
ARR	$576/2223 - 400/2221 = 0.079$	
Number needed to treat (1/ARR)	$1/0.079 \approx 13$	
Mean gain in event-free survival	$4.48 - 4.32 = 0.16 \text{ y} (\approx 2 \text{ mo})$	
Proportion with adverse event in placebo group	$\approx 576/2223 (\approx 0.25 \text{ or } 1 \text{ of } 4)$	
Gain in event-free survival distributed among the proportion of persons with adverse events in the placebo group	$2 \text{ mo}/0.25 = 8 \text{ mo}$	

* Estimations are based on the 4S study (16). 4S = Scandinavian Simvastatin Survival Study; ARR = absolute risk reduction.

Appendix Table 2. Estimation of Corresponding Magnitudes of Number Needed to Treat and Postponements of Adverse Events after 5 Years of Drug Therapy in the Hip Fracture Group*

Variable	Vertebral Fracture Group		Clinical Fracture Group	
	Placebo	Intervention	Placebo	Intervention
Participants, <i>n</i>	1005	1022	812	819
Annual incidence of hip fractures	0.0077	0.0037	0.0053	0.0023
Our estimations				
Fractures in placebo groups, <i>n</i>	59			
Fractures in intervention groups, <i>n</i>	28			
ARR	$59/(1005 + 812) - 28/(1022 + 819) = 0.0173$			
Number needed to treat (1/ARR)	$1/0.0173 \approx 57$			
Fracture-free life-years in placebo groups	8935.7			
Fracture-free life-years in intervention groups	9015.4			
Fracture-free life-years gained	$9015.4 - 8935.7 = 79.7$			
Mean gain in fracture-free survival if the placebo group received the intervention	$79.7/(1005 + 812) = 0.044 \text{ y} (\approx 16 \text{ d})$			
Proportion with hip fractures in the placebo group	$59/(1005 + 812) \approx 0.032 \approx 3 \text{ of } 100$			
Gain in fracture-free survival distributed among the proportion of persons with hip fractures in the placebo group	$0.044 \text{ y}/0.032 (\approx 16 \text{ mo})$			

* Estimations are based on the FIT (17). The FIT had 2 study groups: the vertebral fracture group, which included women who had vertebral fractures identified on radiographs at baseline, and the clinical fracture group, which included women without vertebral fracture but who had a femoral neck T-score less than -1.6 at baseline. We used the annual incidence of hip fractures in the different groups to calculate fracture-free life-years gained after 5 years of therapy. For each of the 5 years, we assumed that fractures occurred in the middle of the year. ARR = absolute risk reduction; FIT = Fracture Intervention Trial.