

## Therapeutic decisions by number needed to treat and survival gains:

a cross-sectional survey of lipid-lowering drug recommendations

### Abstract

#### Background

Previous studies suggest that lay people have difficulties with evaluating effect size in terms of number needed to treat (NNT), but they are sensitive to effect size in terms of survival gains.

#### Aim

To explore whether GPs and internists are sensitive to NNT and survival gains when considering a lipid-lowering drug therapy.

#### Design and setting

Cross-sectional survey of primary prevention of cardiovascular disease with random allocation to different scenarios.

#### Method

GPs ( $n = 450$ ) and internists ( $n = 450$ ) were posted a vignette presenting a high-risk patient and a novel drug, 'neostatin'. The benefit was described in terms of NNT or mean gain in disease-free survival. Each physician was randomly allocated to one version of the vignette. Outcome measures were evaluation of 'neostatin' on a Likert scale (0: very poor choice, 10: very good choice) and the proportion recommending 'neostatin'.

#### Results

A total of 477 responses (53%) were received. Among responders to NNT scenarios, 26%, 31%, and 43% recommended 'neostatin' for NNT values of 34, 17, and 9 respectively. With equivalent disease-free survival gains of 9, 17, and 32 months, 40%, 49%, and 52% respectively recommended the drug. On the rating scale, mean values were 4.7, 5.0, and 5.5 across the respective NNT scenarios and 5.2, 6.2, and 6.1 across the scenarios presenting survival gains. Differences in trends between the two formats were not statistically significant. In total, 33% recommended 'neostatin' when presented with NNT values, compared to 47% when presented with survival gain ( $\chi^2 = 9.2$ ,  $P = 0.002$ ).

#### Conclusion

Physicians presented with survival gains were more likely to recommend the therapy than those presented with NNT. Sensitivity to effect size was similar for both effect formats.

#### Keywords

cardiovascular diseases; decision making; health communication; primary prevention.

### INTRODUCTION

Medical doctors frequently encounter patients for whom long-term risk-reducing drug therapies might be considered. These are complex decisions where therapeutic effectiveness, costs, patient preferences, and probabilities of relevant outcomes should count.<sup>1</sup> However, knowledge about how doctors and their patients actually make such decisions is scant.<sup>2</sup>

With respect to benefits, traditional measures are absolute risk reduction (ARR), relative risk reduction (RRR), and number needed to treat (NNT). Alternatively, to account for the time dimension, benefits may be described in terms of gain in disease-free survival.<sup>3</sup> In terms of survival curves, the first kind of effect measure usually corresponds to the vertical distance between the intervention and control groups at a fixed point in time (Figure 1). Aggregated benefits over time are represented by the area between the two survival curves, which measure the gain in (disease-free) life expectancy. It is noteworthy that towards the end of life, the vertical distance between the curves (that is, ARR), by necessity, decreases towards zero (Figure 1). As the NNT is the inverse value of ARR, it will hence increase towards infinity. These effect measures may therefore have limited ability to convey the benefits of lifelong therapies. On the other hand, the

area between the curves continues to increase over the entire life span (Figure 1), indicating that the longer the patients receive treatment, the greater the benefit in terms of gain in life expectancy.

Presenting treatment benefits in different formats (for example, ARR versus RRR,<sup>4</sup> gains versus losses,<sup>5</sup> survival versus mortality, or numeric versus verbal or visual formats<sup>6</sup>) may influence the decisions of physicians as well as patients. In other words, decisions may be subject to framing effects. A consistent finding is that decision makers are more inclined to accept a therapy when the benefit is explained in terms of RRR rather than ARR or NNT.<sup>4,5</sup> In general, however, framing effects are heterogeneous and sensitive to the exact wording used to convey the treatment effects,<sup>4</sup> and there is no consensus that a single format is superior to others in communicating risk reductions.<sup>7</sup> Empirical evidence suggests that lay people, when considering risk-reducing interventions, are insensitive to effect size in terms of NNT<sup>8-10</sup> and RRR,<sup>11</sup> possibly due to innumeracy<sup>12</sup> or lack of experience with evaluating such numbers.<sup>13,14</sup> On the other hand, lay people are sensitive to the magnitude of gain in disease-free life expectancy explained as the extent to which hip fractures<sup>10</sup> and heart attacks<sup>15</sup> are postponed. In one study, however, it was observed that lay people

**PA Halvorsen**, MD, PhD, researcher, Research Unit for General Practice; research director, National Centre of Rural Medicine, Institute of Community Medicine, University of Tromsø, Tromsø, Norway. **TF Wisløff**, MSc, senior researcher, Norwegian Knowledge Centre for Health Services, Oslo, Norway. **H Støvring**, MSc, PhD, associate professor, Department of Biostatistics, School of Public Health, Aarhus University, Aarhus, Denmark. **OG Aasland**, MD, MHA, director, The Research Institute, The Norwegian Medical Association, Oslo, Norway and Department of Health Management; professor Health Economics, University of Oslo, Oslo, Norway. **IS Kristiansen**, MD, MPH, PhD, professor,

Department of Health Management and Health Economics, University of Oslo, Oslo, Norway.

#### Address for correspondence

Peder A Halvorsen, Svartaksveien 15, 9516 Alta, N-Norway.

**E-mail:** peder.halvorsen@krafslagget.no

**Submitted:** 29 September 2010; **Editor's response:** 22 October 2010; **final acceptance:** 31 January 2011.

#### ©British Journal of General Practice

This is the full-length article (published online 1 Aug 2011) of an abridged version published in print. Cite this article as: **Br J Gen Pract 2011; DOI: 10.3399/bjgp11X588448.**

## How this fits in

Previous studies suggest that lay people, when considering risk-reducing drug therapies, have difficulties with evaluating effect size in terms of number needed to treat (NNT). On the other hand, they are sensitive to effect size in terms of survival gains. In the present study, medical doctors were about equally sensitive to NNTs and survival gains, but physicians presented with survival gains were more likely to recommend a drug therapy than those presented with NNT values. Physicians should perhaps have easy access to both kinds of effect measures.

were more inclined to accept a hypothetical preventive drug therapy when informed about NNTs rather than equivalent gains in disease-free life expectancy.<sup>14</sup>

A powerful determinant of patients' decisions is their doctor's opinion.<sup>16,17</sup> Consequently, whether physicians are sensitive to effect size in their recommendations might influence decisions about drug therapies. Previously, it was shown that physicians were sensitive to the magnitude of NNT when considering a hypothetical drug aimed at prevention of cardiovascular death.<sup>18</sup> In this study, however, a case description with a short time frame (5 years) and few clinical details was used, and two randomly chosen NNTs (50 versus 200) were compared, all of which may limit the resemblance to real-life decisions.

Given that lay people responded differently to NNTs and survival gains,<sup>14</sup> this study aimed to explore whether the same is true for physicians. The present study tested the hypotheses that physicians, when

considering a lipid-lowering drug therapy, are sensitive to the magnitude of NNT as well as gain in disease-free life expectancy. Furthermore, it tested whether physicians would make different judgements of the drug therapy when presented with NNTs rather than equivalent survival gains. The present study, like the previous one,<sup>14</sup> was based on clinical vignettes, but an effort was made to mimic real-life decisions more closely using richer case descriptions, extended time frames, and effect sizes pertinent to a commonly used statin, simvastatin.

## METHOD

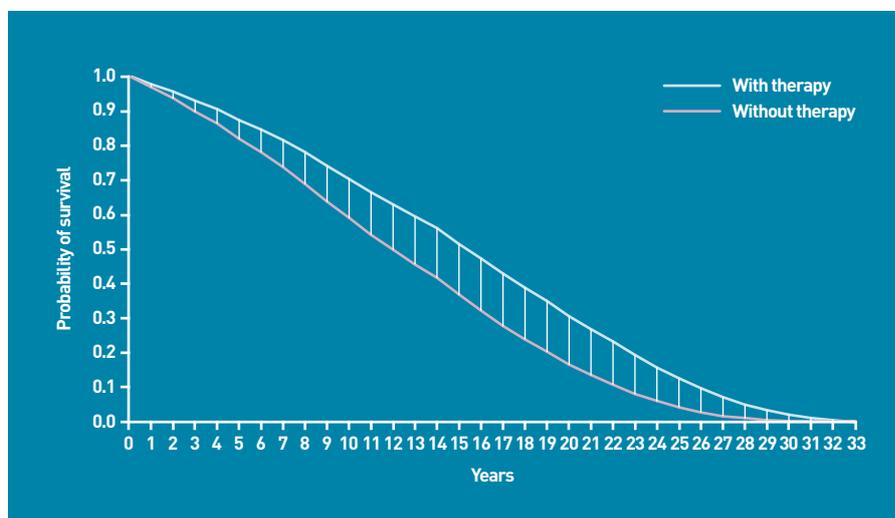
### Trial design and participants

The study was a cross-sectional survey of physicians with random allocation to different versions of a clinical vignette. In January 2009 two random samples of GPs ( $n=450$ ) and internists (specialists in internal medicine,  $n=450$ ) were drawn from the member registry of the Norwegian Medical Association.

### Intervention

Participants were posted a questionnaire that presented a 55-year-old male patient seeking cardiovascular risk assessment due to a family history of premature coronary heart disease (Table 1). The case description included information about smoking and exercise habits, blood pressure, lipids, glucose, body mass index, waist-hip ratio, and electrocardiographic findings. An estimate of the 10-year absolute risk of ischaemic heart disease was also provided (Table 1).<sup>19</sup> Prior to the present study, the case description was piloted among 111 GPs, revised, and tested again among GPs and internists ( $n=218$  and 207 respectively). The proportion

**Figure 1. Hypothetical survival curves illustrating estimation of different effect measures. Vertical lines: differences between proportions at fixed points in time; that is, absolute risk reduction (ARR). Number needed to treat is the inverse value of ARR; that is,  $1/ARR$ . Area between curves: gain in (event-free) life expectancy.**



**Table 1. Characteristics of a 55-year-old male patient seeking cardiovascular risk assessment due to a family history of premature coronary heart disease**

	Details
Family history of coronary heart disease	Father died at 52 years
Smoking	Quit 10 years ago
Physical fitness	Excellent
Total cholesterol	6.5 mmol/l
High-density lipoprotein cholesterol	1.1 mmol/l
Low-density lipoprotein cholesterol	4.5 mmol/l
Triglycerides	2.0 mmol/l
Glucose	4.3 mmol/l
Systolic blood pressure	158 mmHg
Diastolic blood pressure	96 mmHg
Body mass index	24.5 kg/m <sup>2</sup>
Waist-hip ratio	1.0
Electrocardiogram	Unremarkable
10-year risk of ischaemic heart disease <sup>a</sup>	20-30 out of 100

<sup>a</sup>A risk chart from a Norwegian guideline for cardiovascular disease prevention was used to calculate the 10-year risk of ischaemic heart disease.<sup>19</sup> The vignette was presented to 450 internists and 450 GPs.

indicating that the case description was representative of patients they might encounter in their own practice increased from 78% to 88%.

After the case description, a hypothetical lipid-lowering drug labelled 'neostatin' was presented. The drug was to be taken once daily for the rest of the patient's life. Costs and side-effects were described as being similar to other statins. For participants receiving NNT information, the benefit of neostatin was described in terms of the NNT to prevent one incident of cardiovascular disease, such as angina pectoris, heart attack, heart failure, stroke, or cardiovascular death. For participants receiving survival gain information, the benefit of neostatin was described as the average gain in life expectancy free of the same cardiovascular outcomes.

A validated statistical model of cardiovascular disease prevention in Norway (the NorCaD model<sup>20</sup>) was used to derive NNTs and survival gains. First, the model was run using the risk profile of the patient and RRRs established for simvastatin to derive a corresponding NNT and survival gain. Based on a systematic review,<sup>15</sup> RRRs were set at 7%, 8%, 15%, 17%, and 23% for cardiovascular death, heart failure, angina pectoris, stroke, and heart attack respectively. Then the model was run using simvastatin RRRs multiplied by 0.5 and 2.0 respectively to yield another two pairs of NNT and survival gain values.

In each case, lifelong therapy was modelled; however, the notion of 'NNT for lifelong therapy' is tricky. This is because towards the end of life, ARR must approach

zero, and hence the NNT increases to infinity. Therefore, the study used the point in time where NNT was at its 'best'; that is, where the vertical distance between the modelled survival curves was greatest for this particular patient. This time point was identified as occurring after 24 years of therapy. Thus, the procedure yielded three NNTs after 24 years of therapy (9, 17, and 34) and three gains in disease-free survival (9, 17, and 32 months) respectively. These figures were used in six different versions of the questionnaire.

### Randomisation

GPs and internists were randomly allocated to receive only one of the six versions by means of a computerised random number generator (Figure 2). Physicians were informed that they were allocated to different versions of the questionnaire, but not about how the questionnaires differed.

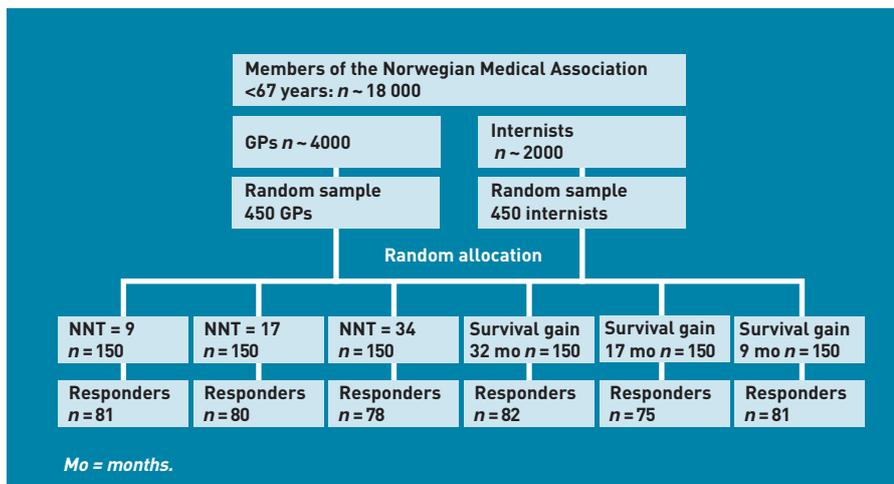
### Outcomes

Responders were asked to rate 'neostatin' on a Likert scale anchored at 0 ('a very poor choice') and 10 ('a very good choice'). They were also asked whether they would recommend 'neostatin' for the patient. Possible response categories were certainly, probably, probably not, and certainly not. Ratings of 'neostatin' and proportions recommending 'neostatin' were the primary outcome measures. Physicians' responses were dichotomised so that those who answered 'certainly' or 'probably' were counted as recommending the therapy, whereas those who responded otherwise were not.

Effect format (that is, NNT or survival gain) and level were the primary independent variables. The physicians' age, sex, and specialty (internist versus GP) were included as secondary independent variables and possible effect modifiers.

### Sample size

The study was powered to detect a 1.5 point difference in mean Likert scores between NNT and disease-free survival gain at each effect level using the standard deviation (SD) from the pilot study (SD = 2.63), power of 0.8, and significance level of 0.05. For the mid effect level, it was hypothesised that about 85% and 60% might recommend 'neostatin' in the NNT and survival-gain scenarios respectively. To detect these differences would require about 50 and 55 responders in each group for mean ratings and proportions, respectively. A computer simulation based on 2500 datasets using ordinary least squares and logistic



**Figure 2. Study flow diagram.** GPs and internists were randomly assigned to clinical vignettes presenting the benefit of a hypothetical lipid-lowering drug therapy, 'neostatin', in terms of number needed to treat (NNT) to prevent one incident of cardiovascular disease after 24 years of therapy, or gain in disease-free survival on lifelong therapy.

regression analyses indicated that 55 responders in each group would yield at least 97% and 74% power to detect a 1.0 difference in mean rating and a 12% difference in proportions per effect level respectively. Based on the worst case response rates in the pilot studies (about 30%), it was decided to invite 150 physicians in each group.

### Statistical methods

Analysis used  $\chi^2$  tests and Mann-Whitney rank sum tests to evaluate differences between proportions and ratings respectively. To test for a linear trend in the rating of 'neostatin' across effect levels, ordinary least squares regression with robust variances was used to account for non-normal distributions of outcomes. Trends in the proportions recommending 'neostatin' were tested using logistic regression. The study tested for first-order interactions between effect measures and age, sex, specialty, and effect level, and between effect level and format, by adding product terms to the regression models. Stata software (version 9.2) was used for data analysis. *P*-values below 0.05 were regarded as statistically significant.

### RESULTS

Of 900 invited physicians, 477 (53%)

responded. The six randomised groups were fairly balanced with respect to age, sex, and working position (Table 2). Among physicians who were presented with gain in disease-free survival information, 112 of 238 (47%) would recommend 'neostatin' compared to 80 of 239 (33%) physicians presented with NNT [difference 14% [95% confidence interval (CI) = 5% to 22%],  $\chi^2 = 9.2$  *P* = 0.002]. Similarly, mean ratings of 'neostatin' were 5.9 and 5.1 among physicians presented with survival gain and NNT respectively [difference 0.8 [95% CI = 0.3 to 1.3],  $|z| = 3.2$ , *P* = 0.001 using the Mann-Whitney rank sum test].

At each level of effect size, the proportions recommending 'neostatin' and mean ratings were consistently higher among physicians presented with survival gains than those who were given NNT values, but the differences reached statistical significance for the mid effect level only (Table 3). Across the three levels of effect size, mean ratings and proportions recommending 'neostatin' tended to increase with increasing survival gains and decreasing NNTs (Table 3). There were no statistically significant interactions between effect level and format; that is, no differences in trends across the three levels between the two effect formats (Table 3). However, the study was not powered to detect such differences. Regression analysis indicated that, adjusted for age and sex, internists were more likely than GPs to recommend 'neostatin' (odds ratio [OR] = 1.47, 95% CI = 1.01 to 2.13), and they rated 'neostatin' more highly [mean difference 0.8, 95% CI = 0.3 to 1.2]. There were no statistically significant interactions between type or level of effect measure and age, sex, or specialty.

### DISCUSSION

#### Summary

In this survey, physicians were more likely to recommend a lifelong lipid-lowering drug therapy, and they rated the drug more highly on a Likert scale, when the benefit was explained in terms of gain in disease-free survival rather than the corresponding NNT. In their recommendations and Likert scale ratings, the physicians were about equally sensitive to effect size in terms of NNT and survival gains.

#### Strengths and limitations

The randomised design, which ensured that all physicians were exposed to the same clinical scenario except for the different effect formats, and the careful piloting, are the main strengths of this study. However, there are several limitations. The response

**Table 2. Characteristics of responders in the six study arms<sup>a</sup>**

Scenario	Number needed to treat			Survival gain, months		
	9	17	34	9	17	32
Participants, n	81	80	78	81	75	82
Mean age, years (SD)	53 (10.7)	54 (8.9)	54 (9.1)	51 (8.4)	55 (9.0)	53 (9.0)
Women, %	19	24	21	32	24	20
Working as GP, %	51	49	54	51	51	51

<sup>a</sup>Participants were randomised to one of six vignettes portraying a high-risk individual and a risk-reducing drug therapy. Vignettes differed by effect size and format (NNT or gain in disease free survival).

**Table 3. Physicians' assessment of a hypothetical drug ('neostatin') by effect formats (number needed to treat [NNT] versus survival gain) and effect level**

	Survival gain scenarios	NNT scenarios	Both scenarios combined	Differences
	Proportions recommending 'neostatin', % [95% CI]			% [95% CI] $\chi^2$ statistics
Level 1, NNT = 34, survival gain = 9 months	39.5 [28.7 to 50.3]	25.6 [15.8 to 35.4]	32.7 [25.4 to 40.0]	13.9 [-0.5 to 28.3] $\chi^2 = 3.5, P = 0.062$
Level 2, NNT = 17, survival gain = 17 months	49.3 [37.9 to 60.8]	31.3 [21.0 to 41.5]	40.0 [32.2 to 47.8]	18.1 [2.9 to 33.3] $\chi^2 = 5.3, P = 0.022$
Level 3, NNT = 9, survival gain = 32 months	52.4 [41.5 to 63.3]	43.2 [32.3 to 54.1]	47.9 [40.1 to 55.6]	9.2 [-6.0 to 24.5] $\chi^2 = 1.4, P = 0.238$
<b>Trend analysis<sup>a</sup></b>				
Odds ratio per level [95% CI]	1.3 [1.0 to 1.8]	1.5 [1.1 to 2.1]		
Odds ratio for interaction term format × level [95% CI]		1.2 [0.7 to 1.8]		
Evaluation of 'neostatin' on a 0–10 scale, <sup>b</sup> score [95% CI]				Mean [95% CI], rank sum statistics
Level 1, NNT = 34, survival gain = 9 months	5.2 [4.6 to 5.8]	4.7 [4.1 to 5.3]	5.0 [4.6 to 5.4]	0.5 [-0.3 to 1.4] $ z  = 1.3, P = 0.189^c$
Level 2, NNT = 17, survival gain = 17 months	6.2 [5.7 to 6.8]	5.0 [4.4 to 5.6]	5.6 [5.2 to 6.0]	1.2 [0.4 to 2.0] $ z  = 3.1, P = 0.002^c$
Level 3, NNT = 9, survival gain = 32 months	6.1 [5.6 to 6.6]	5.5 [4.9 to 6.2]	5.8 [5.4 to 6.2]	0.6 [-0.3 to 1.4] $ z  = 1.1, P = 0.264^c$
<b>Trend analysis<sup>d</sup></b>				
Regression coefficient per level	0.42 [0.03 to 0.81]	0.41 [-0.03 to 0.85]		
Regression coefficient for interaction term format × level [95% CI]		-0.01 [-0.59 to 0.58]		

<sup>a</sup>Logistic regression analysis of trend across the three levels of effect size. <sup>b</sup>0 = 'a very poor choice', 10 = 'a very good choice'. <sup>c</sup>Mann-Whitney rank sum test. <sup>d</sup>Ordinary least squares regression analysis with robust variances.

rate was modest, and it is not known whether the findings are representative for GPs and internists in general. In the vignette, only one risk profile was used; therefore, the study findings may not apply to other patients with other risk profiles. Furthermore, the vignette did not include measures of uncertainty around the effect estimates, which may limit their validity. The two pilot studies indicated that the face validity of the scenario was fair. There is no guarantee that the responses to vignettes are representative of real-life clinical decisions,<sup>21</sup> but previous studies indicate that vignette techniques may perform reasonably well when validated against other methods.<sup>22,23</sup>

An important issue is whether testing physicians' responses to NNTs and survival gains is at all meaningful. From a practical point of view, these effect measures are often not readily available at the point of care, which makes it difficult for busy clinicians to include them in conversations with patients. Furthermore, given the conventional time frame of 10 years for cardiovascular risk estimates prevailing in clinical guidelines, it may seem artificial to test physicians' responses to survival gains, which they rarely encounter in the scientific literature or elsewhere. On the other hand, when preventive drug therapies such as

lipid-lowering drugs and antihypertensives are recommended for the rest of the patient's life, conveying the benefits accrued over the same time frame (that is, lifelong) seems reasonable.

Interestingly, Trewby *et al* showed that at least half of patients would expect at least 12 months' average prolongation of life to make a hypothetical cholesterol-lowering drug therapy worthwhile.<sup>24</sup> This suggests that patients may find conversations about survival gains meaningful. In a recent analysis, Buyx *et al* suggest using improvements in life expectancy as a criterion for priority setting.<sup>25</sup>

#### Comparison with existing literature

Why did the survival-gain format evoke a more positive attitude to the hypothetical drug therapy than the corresponding NNT? In a previous study, the authors showed that a substantial proportion of physicians interpreted the NNT as a direct measure of the likelihood of benefit from therapy; that is, that only one out of the NNT patients would benefit.<sup>18</sup> This interpretation, which implies that only a small minority of patients will benefit from therapy unless the NNT is very small (for example, less than 5), is probably due to the fact that the NNT reflects the state of affairs measured at a specific point in time.<sup>8</sup> As a tool to convey

benefits accruing over time, the NNT may therefore suffer from limitations that are not inherent in the survival gain format.

A limitation of the survival gain format is that it is not possible to infer the distribution of survival gain among patients. Some may gain a substantial amount of extra time free of disease, while others gain little or nothing.<sup>8,26</sup>

In a previous study, lay people were more inclined to accept a hypothetical therapy when the benefit was explained in terms of NNT rather than survival gain.<sup>14</sup> For example, 93% consented when the NNT to prevent a heart attack was 13, 69% consented when the drug postponed the heart attack by 2 months for all patients, whereas 82% consented when the benefit was postponement of heart attacks by 8 months for one out of four patients. In another study arm, similar findings were reported when participants were presented with a hypothetical drug to prevent hip fractures. These findings are opposite to those of the present study. The time frame in the previous study, however, was only 5 years, as opposed to the lifelong therapy of the present study. Therefore, the survival gains were smaller. To the authors' knowledge, no previous study has tested the effect of presenting NNTs against survival gains to physicians.

From a normative viewpoint, the magnitude of benefit should have weight in decisions regarding drug therapies,<sup>1</sup> but may in practice carry little weight. Lay people were insensitive to effect size in terms of RRR,<sup>11</sup> as well as NNT<sup>8,9</sup> in studies

similar to the present one. However, two studies suggest that lay people are sensitive to effect size in terms of survival gains.<sup>10,15</sup> In the present study, physicians tended to be sensitive to NNT as well as survival gains, although the trends did not always reach statistical significance, possibly due to the limited power to detect such trends. The discrepancy in sensitivity to NNTs between lay people and physicians may be explained by the evaluability hypothesis,<sup>13</sup> which predicts that factors that are difficult to evaluate will bear little weight on people's decisions. It is conceivable that lay people find it more difficult than physicians to evaluate the magnitude of NNT.

#### **Implication for practice and research**

GPs and internists were more inclined to recommend a hypothetical lipid-lowering drug therapy when benefits were explained in terms of disease-free survival gain rather than NNT. When considering lipid-lowering drug therapies, clinicians should perhaps have ready access to both kinds of effect measures. Qualitative studies might help to clarify whether doctors and patients find such measures helpful in their clinical encounters. If sufficiently powered studies were to confirm that physicians are sensitive to survival gains, this effect format might have potential as 'common language' in these situations. In that case, reporting survival gains in clinical trials, and including such effect measures in clinical guidelines for cardiovascular disease prevention, might be warranted.

---

#### **Funding body**

University of Tromsø, Norway, and the Norwegian Medical Association. The funding bodies had no involvement in the design of the study, data collection and interpretation, or decision to submit the manuscript for publication.

#### **Ethical approval**

The study was approved by the Norwegian Social Science Data Services, which is the Privacy Ombudsman for all Norwegian Universities as well as the Research Institute of the Norwegian Medical Association.

#### **Provenance**

Freely submitted; externally peer reviewed.

#### **Competing interests**

The authors have declared no competing interests.

#### **Acknowledgements**

Thank you to the participants for their time and effort.

#### **Discuss this article**

Contribute and read comments about this article on the Discussion Forum:  
<http://www.rcgp.org.uk/bjgp-discuss>

## REFERENCES

1. Hunink MG, Glasziou P, Siegel J, *et al.* *Decision making in health and medicine: integrating evidence and values. 1st edn.* Cambridge: Cambridge University Press, 2001.
2. Kassirer JP, Wong JB, Kopelman RI. *Learning clinical reasoning. 2nd edn.* Philadelphia, PA: Lippincott Williams and Wilkins, 2009.
3. Tan LB, Murphy R. Shifts in mortality curves: saving or extending lives? *Lancet* 1999; **354(9187)**: 1378–1381.
4. Covey J. A meta-analysis of the effects of presenting treatment benefits in different formats. *Med Decis Making* 2007; **27(5)**: 638–654.
5. Edwards A, Elwyn G, Covey J, *et al.* Presenting risk information — a review of the effects of ‘framing’ and other manipulations on patient outcomes. *J Health Commun* 2001; **6(1)**: 61–82.
6. Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Med Decis Making* 2007; **27(5)**: 696–713.
7. Ghosh AK, Ghosh K. Translating evidence-based information into effective risk communication: current challenges and opportunities. *J Lab Clin Med* 2005; **145(4)**: 171–180.
8. Kristiansen IS, Gyrd-Hansen D, Nexoe J, Nielsen JB. Number needed to treat: easily understood and intuitively meaningful? Theoretical considerations and a randomized trial. *J Clin Epidemiol* 2002; **55(9)**: 888–892.
9. Halvorsen PA, Kristiansen IS. Decisions on drug therapies by numbers needed to treat: a randomized trial. *Arch Intern Med* 2005; **165(10)**: 1140–1146.
10. Christensen PM, Brosen K, Brixen K, *et al.* A randomised trial of laypersons’ perception of the benefit from osteoporosis therapy: number needed to treat versus postponement of hip fracture. *Clin Ther* 2003; **25(10)**: 2575–2585.
11. Sorensen L, Gyrd-Hansen D, Kristiansen IS, *et al.* Laypersons’ understanding of relative risk reductions: randomised cross-sectional study. *BMC Med Inform Decis Mak* 2008; **8**: 31.
12. 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(11)**: 884–892.
13. 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(5)**: 576–590.
14. Halvorsen PA, Selmer R, Kristiansen IS. Different ways to describe the benefits of risk-reducing treatments: a randomized trial. *Ann Intern Med* 2007; **146(12)**: 848–856.
15. Dahl R, Gyrd-Hansen D, Kristiansen I, *et al.* Can postponement of an adverse outcome be used to present risk reductions to a lay audience? A population survey. *BMC Med Inform Decis Mak* 2007; **7(1)**: 8.
16. Mazur DJ, Hickam DH, Mazur MD, Mazur MD. The role of doctor’s opinion in shared decision making: what does shared decision making really mean when considering invasive medical procedures? *Health Expect* 2005; **8(2)**: 97–102.
17. Gwyn R, Elwyn G, Edwards A, Mooney A. The problematic of decision-sharing: deconstructing ‘cholesterol’ in a clinical encounter. *Health Expect* 2003; **6(3)**: 242–254.
18. Halvorsen PA, Kristiansen IS, Aasland OG, Forde 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(3)**: 162–166.
19. Meland E, Ellekjaer H, Gjelsvik B, *et al.* [Pharmacological prevention of cardiovascular diseases in general practice]. *Tidsskr Nor Laegeforen* 2000; **120(22)**: 2643–2647.
20. Wisløff T, Selmer R, Halvorsen S, Kristiansen IS. *Norwegian Cardiovascular Disease model (NorCaD) — a simulation model for estimating health benefits and cost consequences of cardiovascular interventions.* Oslo: Norwegian Knowledge Centre for the Health Services, 2008.
21. Jones TV, Gerrity MS, Earp J. Written case simulations — do they predict physicians behavior. *J Clin Epidemiol* 1990; **43(8)**: 805–815.
22. Peabody JW, Luck J, Glassman P, *et al.* Measuring the quality of physician practice by using clinical vignettes: A prospective validation study. *Ann Intern Med* 2004; **141(10)**: 771–780.
23. Peabody JW, Luck J, Glassman P, *et al.* Comparison of vignettes, standardized patients, and chart abstraction — a prospective validation study of 3 methods for measuring quality. *JAMA* 2000; **283(13)**: 1715–1722.
24. Trewby PN, Reddy AV, Trewby CS, *et al.* Are preventive drugs preventive enough? A study of patients’ expectation of benefit from preventive drugs. *Clin Med* 2002; **2(6)**: 527–533.
25. Buyx AM, Friedrich DR, Schone-Seifert B. Ethics and effectiveness: rationing healthcare by thresholds of minimum effectiveness. *BMJ* 2011; **342**: d54.
26. Henderson R, Keiding N. Individual survival time prediction using statistical models. *J Med Ethics* 2005; **31(12)**: 703–706.