Informed Health Choices

Understanding of key concepts relevant to assessing claims about treatment effects: a survey of Norwegian adults

Oxman AD et al.

Working paper, 1 March 2017

Colophon

Title Understanding of key concepts relevant to assessing claims about treatment effects: a survey of Norwegian adults

Authors Oxman, Andrew David
Austvoll-Dahlgren, Astrid
Garratt, Andrew
Rosenbaum, Sarah

Corresponding author Astrid Austvoll-Dahlgren
Astrid.Austvoll-Dahlgren@fhi.no
Norwegian Institute of Public Health
PO Box 4404, Nydalen
N-0403 Oslo, Norway

Keywords health literacy, numeracy, critical thinking, critical appraisal, informed health choices, evidence-based healthcare, outcome measurement

ISBN 978-82-8082-819-4

URL http://www.informedhealthchoices.org/wp-content uploads/2016/08/CURE-survey-report-2017-03-01.pdf

Citation Oxman AD, Austvoll-Dahlgren A, Garratt A, Rosenbaum S. Understanding of key concepts relevant to assessing claims about treatment effects: a survey of Norwegian adults. IHC Working Paper, 2017.

Date March 2017

Plain language summary

Little is known about the ability of Norwegians - or other populations - to assess claims about the effects of "treatments" (any action intended to improve health) and make informed health choices. We measured Norwegian adults' understanding of key concepts that are relevant to assessing claims about treatment effects in 2005, using nine multiple-choice questions. These questions were precursors to the multiple-choice questions in the Claim Evaluation Tools database.

The questions were based in part on the Medical Data Interpretation test by Schwartz and Woloshin. The development also included a literature review, consultation with experts in the field of evidence-based medicine, and an open-meeting with patients contributed to the development of the nine questions. The nine questions assessed four key concepts relevant to assessing claims about treatment effects and numeracy. The questionnaire also included questions about attitudes and background information about the participants.

Of the 2500 Norwegian adults to whom we sent the questionnaire in 2005, 626 (25.0%) responded to the survey. The average age of respondents was 48 years, about half (52.5%) were female, and about half (52.4%) had at least one year of education beyond secondary school.

The proportion of correct answers to the nine questions varied from only 19.2% for understanding that an outcome may be associated with a treatment, but not caused by the treatment; to 76.5% for a question assessing numeracy (understanding that risk estimates require a denominator). The median proportion of correct scores for all nine questions was 59.3%. The average number of correct answers was 4.92 out of nine (sd=2.06).

Based on these findings, we concluded that there are important gaps in Norwegian adults' understanding of key concepts that they should understand and apply when assessing claims about treatment effects.

Abstract

Background

Little is known about the ability of Norwegians to assess claims about the effects of "treatments" (any action intended to improve health) and make informed health choices. We measured their understanding of key concepts relevant for assessing treatment claims.

Objective

To assesses Norwegian adults' understanding of key concepts relevant to assessing claims about treatment effects.

Methods

The questions were based in part on the Medical Data Interpretation test by Schwartz and Woloshin. The development also included a literature review, consultation with experts in the field of evidence-based medicine, and an open-meeting with patients contributed to the development of the nine questions. We sent a questionnaire with these nine questions to a stratified random sample of 2500 Norwegian adults in 2005.

Results

The nine questions assessed four key concepts relevant to assessing claims about treatment effects, as well as numeracy. Of the 2500 Norwegian adults to whom the questionnaire was sent, 626 (25.0 %) responded to the survey. The mean age of respondents was 48.4 (sd=17.3), 326 (52.5%) were female, and 301 (52.4%) had at least one year of education beyond secondary school. The questionnaire had low levels of missing data and scores were approximately normally distributed with low floor and ceiling effects (12 (1.9%) and 14 (2.2%) respondents respectively). We found expected associations between correct responses to the nine questions and education level, research experience, age, and Internet access.

The proportion of correct answers to the nine questions varied from 19.2% for understanding that an outcome may be associated with a treatment, but not caused by the treatment; to 76.5% for a question assessing numeracy (understanding that risk estimates require a denominator). The median proportion of correct scores for all nine questions was 59.3%. The average number of correct answers was 4.92 out of nine (sd=2.06).

Conclusion

There are important gaps in Norwegian adults' understanding of key concepts that are necessary to understand and apply when assessing claims about the effects of treatments.

Background

Evidence-informed patient choice[1] brings together two approaches to achieving well-informed healthcare choices: evidence-based healthcare[2] and shared decision making.[3] It is rooted in the patient's right to choose to receive a treatment or not, based on their own values and an informed decision. This requires considering whether the desirable effects of a "treatment" (any action intended to improve health) outweigh the undesirable effects.

However, decision-making is increasingly difficult because of the vast amount of information available.[4] If "patients" (anyone confronted with a claim about the effects of use a treatment or a decision about whether to use a treatment) are unable to assess claims about the effects of healthcare, decisions may be made on the wrong grounds. Various initiatives have been devised to improve public understanding of the effects of healthcare, including guidelines for health reports in the mass media, training, and internet resources.[5-11] However, there has been little rigorous evaluation of these initiatives, including the measurement of patients' ability to assess claims about treatment effects.[12] Such measurement is necessary for assessing the effectiveness of educational interventions aimed at improving people's understanding of healthcare research and the purported effects of healthcare, whether in the mass media or elsewhere.[13]

This paper is based on work that we undertook in 2005, in our first attempt at developing questions to measure people's understanding of healthcare research and their ability to assess claims about treatment effects. We developed the "Consumer Understanding of Statistics used in Health Reports about the Effects of Healthcare" (CURE) questionnaire as a brief, self-administered questionnaire intended for use in the evaluation of interventions designed to improve people's understanding of health research and their ability to assess claims about treatment effects. This work was a precursor to subsequent work on the Claim Evaluation Tools database, which includes a battery of multiple-choice questions that can be used to measure people's ability to assess claims about treatment effects and to make informed health choices.[14] We report here the findings of a survey of Norwegian adults undertaken in 2005 using the CURE questionnaire.

Methods

Development of the questions

We reviewed the content of existing questionnaires designed for clinicians and patients for potential relevance.[10,11,16-27] The questions were based in part on the Medical Data Interpretation test by Schwartz and Woloshin. A literature review, consultation with experts in the field of evidence-based medicine, and an open-meeting with patients contributed to the development of the nine questions. Clinical scenarios are frequently used in teaching evidence-based medicine and in evaluating health professionals' evidence-based healthcare skills. We developed similar scenarios relevant to patients for the questions used in the CURE questionnaire.

We sought input from experts in evidence-based healthcare regarding the sensibility of draft questions.[28,29] We also invited patients and representatives of patient organisations to an open meeting to discuss the draft questions, identify unclear or ambiguous questions, and identify important issues that were missing from the questions. Participants were asked to comment on the relevance of issues and comprehensibility of the draft questions.

Based on these inputs, we produced nine questions to measure an individual's ability to understand effectiveness research and assess treatment claims (Box 1). The questionnaire also included questions about participants' attitudes and background (Appendix 1). Following those revisions, we piloted the questionnaire by sending it to a stratified random sample of 667 adults in Norway.

Box 1. Questions measuring participants' ability to assess claims about treatment effects

1. You have been told by your doctor that you have a serious heart disease. Which is the most important factor when deciding to use a drug to treat the disease?

- ☐ The number of deaths it prevents
- ☐ That the benefits outweigh the side-effects
- ☐ The number of people who have the disease
- ☐ Don't know

2. A new study finds that there are 30 deaths among people who eat broccoli regularly compared to 100 deaths among people who don't eat broccoli at all.

Based on this study, which statement best describes how eating broccoli relates to death?

- ☐ Lowers the risk of death
- ☐ Doesn't change the risk of death
- ☐ Raises the risk of death
- ☐ Can't tell from this information

3. About 6,500 Norwegians will be diagnosed with diabetes this year.

What is your best guess about a Norwegian's chance of being diagnosed with diabetes in the next year?

- 6,500 to one ☐
- 6,500 divided by the number of Norwegians ☐
- Don't know ☐

Mrs. Jones is told by her doctor that she has a 28 in 1000 chance of dying from cancer and a 59 in 1000 chance of dying from a stroke.

Mrs Jones' doctor now tells her that a new pill, "ANTI-STROKE", will lower her chance of dying from stroke by 50%. Another pill, "ANTI-CANCER", will lower her chance of dying from cancer by 50%.

4. She can only take 1 pill. Assuming the two pills are equally safe and the cost is the same, which do you suggest she take if her only goal is to lower her chance of dying?

- ANTI-STROKE pill ☐
- ANTI-CANCER pill ☐
- Either pill, it makes no difference ☐
- Don't know ☐

5. Mrs Jones decides to take the ANTI-CANCER pill. Now, what is her chance of dying from cancer?

- 0 in 1000 ☐
- 7 in 1000 ☐
- 14 in 1000 ☐
- 21 in 1000 ☐
- Don't know ☐

6. What do we mean when we say that the result of the study of a treatment is "statistically significant"?

- That the result is important ☐
- That the results are unlikely to be due to chance ☐
- That the results represent a big difference ☐
- Don't know ☐

7. What does a confidence interval tell us about a treatment effect?

- It is the area where the result is unlikely to be due to systematic error ☐
- It is the area where the treatment effect is so big that it should be used in clinical practice ☐
- It is the area where it is most probable that we will find the true treatment effect ☐
- Don't know ☐

8. Which of the following programmes for people who have had a heart attack is most effective?
In the three years after a heart attack:

- Programme A would reduce the rate of deaths by 20% ☐
- Programme B would reduce the number of deaths by 3% ☐
- Programme C would increase patient survival from 84% to 87% ☐
- Programme D would mean that 31 people needed to enter the programme to prevent one death ☐
- They are all equally effective ☐
- Don't know ☐

9. A new study found that people drinking alcohol regularly were more likely to have cancer. Based on this study, which statement best describes how drinking alcohol relates to cancer?

Alcohol lowers the risk of cancer ☐
Alcohol doesn't change the risk of cancer ☐
Alcohol raises the risk of cancer ☐
Can't tell from this information ☐

Correct answers:
1 That the benefits outweigh the side-effects
2 Can't tell from this information
3 6,500 divided by the number of Norwegians
4 ANTI-STROKE pill
5 14 in 1000
6 That the results are unlikely to be due to chance
7 It is the area where it is most probable that we will find the true treatment effect
8 They are all equally effective
9 Can't tell from this information

Participants, sample size, and data collection

Questionnaires were mailed to a stratified random sample of 2500 people in Norway over 18 years old, based on gender and where they lived. The sample size calculation for the survey of Norwegian adults was based on a response rate of 25% in the pilot study. To increase the response rate, we offered gift certificates to respondents through a lottery. An accompanying information letter asking for consent was included together with a stamped and addressed return envelope. One reminder was sent. Questionnaires were anonymous, but response envelopes were given a number so that we could avoid sending reminders to participants who had already returned the questionnaire, and could identify winners of the gift certificates.

The survey was given ethical approval by the Norwegian Social Science Data Services.

Statistical analysis

Responses to the questions were summed and the total scores were assessed for floor and ceiling effects. Such end effects are undesirable because they will limit the questionnaires ability to measure cross-sectional and longitudinal score differences.

Questionnaire scores were assessed for construct validity through comparisons with education and age. It was hypothesised that those with more years of education would be more familiar with the concepts, and hence have higher scores. It was further hypothesised that health professionals and people with research experience would have higher scores. The differences were tested for both individual questions and total scores. For binary variables, Pearson's Chi-Squared test and the t-test were used to compare differences in responses to individual

questions and total scores respectively. For ordinal variables, Spearman's correlation coefficient was used.

Stepwise multiple regression analysis was used to assess which variables made the strongest contribution to explaining variation in total scores. The total scores were the dependent variable. Age, gender, education, whether the respondent was a health professional, research experience, prior participation in clinical research, and internet access were the independent variables.

Results

Of the 2500 Norwegian adults to whom we sent a questionnaire, 626 (25.0%) returned a completed questionnaire. Of these, 326 (52.5%) were female, their mean age was 48.4 (sd=17.3) years, and 301 (52.4%) had at least one year of education beyond secondary school (Table 1).

There were very few missing responses. The highest proportion of missing responses was 1.4% for the sixth question, which addressed the meaning of statistical significance. The lowest and highest number of correct responses were for questions 9 (19.2%) and 3 (76.5%) (Figure 1). The average number of correct answers was 4.92 (sd=2.06) out of 9, and the scores were approximately normally distributed with low floor and ceiling effects of 12 (1.9%) and 14 (2.2%) respondents respectively. The number of correct responses for each question and the total scores by sociodemographic variables are shown in Table 1.

Respondents with the lowest education level had the lowest proportion of correct responses for all questions. Those with a PhD had the highest proportion of correct responses for all but question 7. There was a correlation between education level and the number of correct responses for eight of the nine questions and for the total scores, which were unlikely to have occurred by chance ($P<0.01$ for six questions and the total score, and $P<0.05$ for two questions). Respondents with research experience had a higher proportion of correct responses for all questions, and these differences were unlikely to have occurred by chance for five of the nine questions and the total scores ($P<0.01$ for three questions and the total score and $P<0.05$ for two questions). Health professionals and non-health professionals had similar proportions of correct responses for all nine questions (Figure 2).

Respondents with Internet access had a higher proportion of correct responses for eight questions. This was unlikely to have occurred by chance ($P<0.05$) for

seven questions and the total scores ($P<0.01$ for six questions and the total score and $P<0.05$ for one question). Respondents who had participated in clinical research had a higher proportion of correct responses for eight of the nine questions, compared to respondents who had not participated in clinical research. However, this was unlikely to have occurred by chance for just one question and for the total scores ($P<0.05$).

The proportion of correct responses declined with age for eight questions. The only exception was question 9, for which those under 50 years old scored lower than those who were 50 years old or older. The correlation between age and the number of correct responses was unlikely to have occurred by chance for five questions and for the total scores ($P<0.01$). The proportion of correct answers for each question and the total scores were similar for women and men.

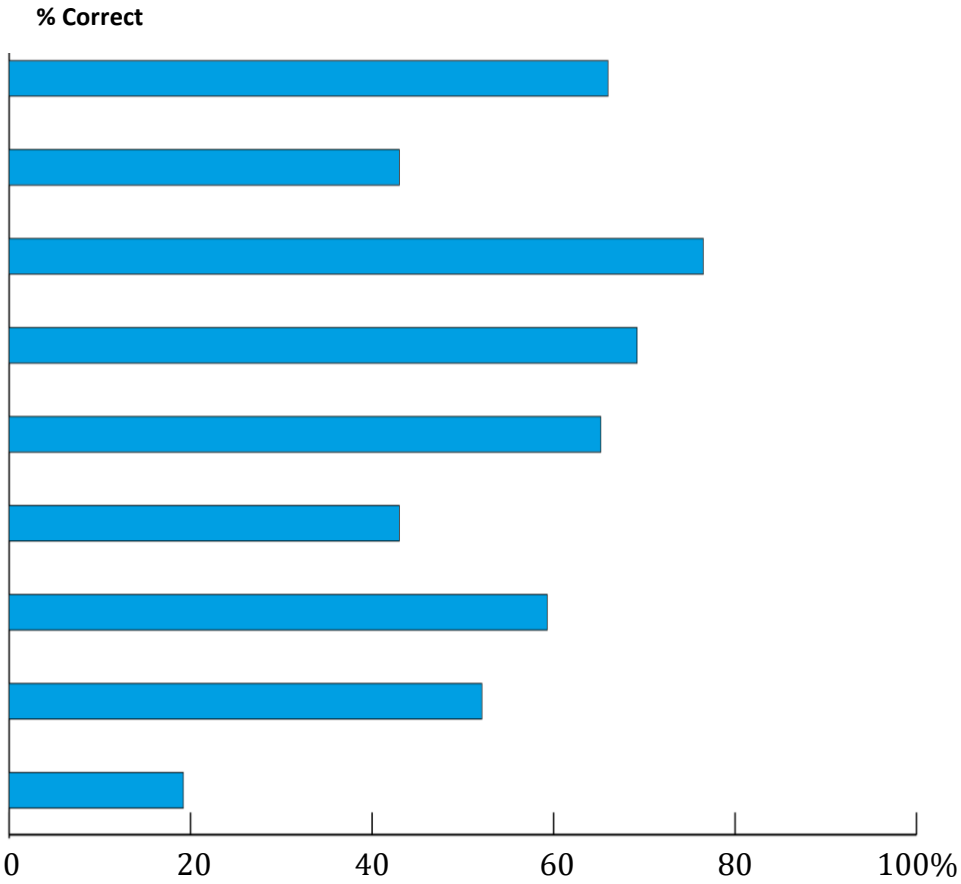
Table 1. Numbers (%) correct answers and average (standard deviation) total scores

Variable (n) (n=626)	1	2	3	4	5	6	7	8	9	Total scores
Age:										
18-34 (142)	97 (68.3%)	78 (54.9%)**	121 (85.2%)**	115 (81.0%)**	108 (76.1%)**	66 (46.5%)**	76 (53.5%)	83 (58.5%)	21 (14.8%)	5.39 (1.81)**
35-49 (189)	131 (69.3%)	89 (47.1%)	153 (81.0%)	141 (74.6%)	135 (71.4%)	98 (51.9%)	127 (67.2%)	93 (49.2%)	33 (17.5%)	5.29 (1.93)
50-64 (163)	105 (64.4%)	65 (39.9%)	128 (78.5%)	110 (67.5%)	103 (63.2%)	65 (39.9%)	96 (58.9%)	92 (56.4%)	33 (20.3%)	4.89 (2.09)
65-79 (99)	63 (63.6%)	27 (27.3%)	56 (56.6%)	53 (53.5%)	50 (50.5%)	35 (35.4%)	53 (53.5%)	41 (41.4%)	28 (28.3%)	4.10 (2.17)
80-89 (22)	11 (50.0%)	6 (27.3%)	11 (50.0%)	10 (45.5%)	8 (36.4%)	3 (13.6%)	5 (22.7%)	13 (59.1%)	2 (9.1%)	3.14 (1.52)
Gender:										
Female (326)	217 (66.6%)	147 (45.1%)	244 (74.9%)	224 (68.7%)	212 (65.0%)	131 (40.2%)	192 (58.9%)	171 (52.5%)	63 (19.3%)	4.91 (1.9%)
Male (295)	194 (65.8%)	120 (40.7%)	229 (77.6%)	207 (70.2%)	195 (66.1%)	137 (46.4%)	170 (57.6%)	154 (52.2%)	55 (18.6%)	4.95 (2.2%)
Education:										
≤ secondary school (301)	187 (62.1%)*	94 (31.2%)**	193 (64.1%)**	169 (56.2%)**	154 (51.2%)**	52 (17.3%)**	161 (53.5%)	132 (43.9%)**	48 (16.0%)*	3.95 (1.80)**
1-3 years university (148)	104 (70.3%)	61 (41.2%)	123 (83.1%)	119 (80.4%)	109 (73.7%)	91 (61.5%)	98 (66.2%)	78 (52.7%)	26 (17.6%)	5.47 (1.83)
Masters level (161)	110 (68.3%)	106 (65.8%)	148 (91.9%)	133 (82.6%)	136 (84.5%)	116 (72.1%)	98 (60.9%)	106 (65.8%)	38 (23.6%)	6.16 (1.68)
PhD (9)	9 (100%)	6 (66.7%)	9 (100%)	9 (100%)	8 (88.9%)	9 (100%)	5 (55.6%)	8 (88.9%)	4 (44.4%)	7.44 (1.59)
Health professional:										
No (493)	331 (67.1%)	205 (41.6%)	369 (74.9%)	337 (68.4%)	326 (66.1%)	207 (42.0%)	284 (57.6%)	256 (51.9%)	90 (18.3%)	4.88 (2.07)
Yes (124)	77 (62.1%)	60 (48.4%)	100 (80.7%)	92 (74.2%)	78 (62.9%)	60 (48.4%)	76 (61.3%)	67 (54.0%)	27 (21.8%)	5.14 (1.94)
Research experience:										
No (553)	361 (65.3%)	224 (40.5%)**	414 (74.9%)*	377 (68.2%)	354 (64.0%)*	217 (39.2%)**	317 (57.3%)	273 (47.4%)**	101 (18.3%)	4.77 (1.98)**
Yes (68)	50 (73.5%)	43 (63.2%)	59 (86.8%)	54 (79.4%)	53 (77.9%)	51 (75.0%)	45 (66.2%)	52 (76.5%)	17 (25.0%)	6.24 (2.02)
Internet access:										
No (105)	57 (54.3%)**	25 (23.8%)**	56 (53.3%)**	47 (44.8%)**	44 (41.9%)**	17 (16.2%)**	52 (49.5%)*	52 (49.5%)	25 (23.8%)	3.57 (1.81)**
Yes (515)	353 (68.5%)	242 (47.0%)	416 (80.8%)	383 (74.4%)	362 (70.3%)	251 (48.7%)	310 (60.2%)	273 (53.0%)	93 (18.1%)	5.21 (1.97)
Asked to take part in clinical research:										
No (510)	337 (66.1%)	206 (40.4%)**	383 (75.1%)	350 (68.6%)	326 (63.9%)	209 (41.0%)*	287 (56.3%)*	263 (51.6%)	102 (20.0%)	4.83 (2.04)**
Yes (109)	72 (66.1%)	61 (56.0%)	89 (81.7%)	80 (73.4%)	80 (73.4%)	58 (53.2%)	74 (67.9%)	60 (55.1%)	16 (14.7%)	5.41 (1.95)
Taken part in clinical research:										
No (538)	353 (65.6%)	218 (40.5%)**	409 (76.0%)	372 (69.1%)	347 (64.5%)	227 (42.2%)	310 (57.6%)	274 (60.9%)	105 (19.5%)	4.86 (2.03)*
Yes (82)	57 (69.5%)	49 (59.8%)	64 (78.1%)	59 (72.0%)	60 (73.2%)	41 (50.0%)	52 (63.4%)	50 (61.0%)	13 (15.9%)	5.43 (2.05)

For binary variables Pearson's Chi-Squared test and the t-test were used for individual questions and total scores respectively. For ordinal variables, Spearman's correlation was used. One asterisk (*) indicates $P<0.05$ and two asterisks (**) indicates $P<0.01$.

Figure 1. Proportion of correct answers for each question

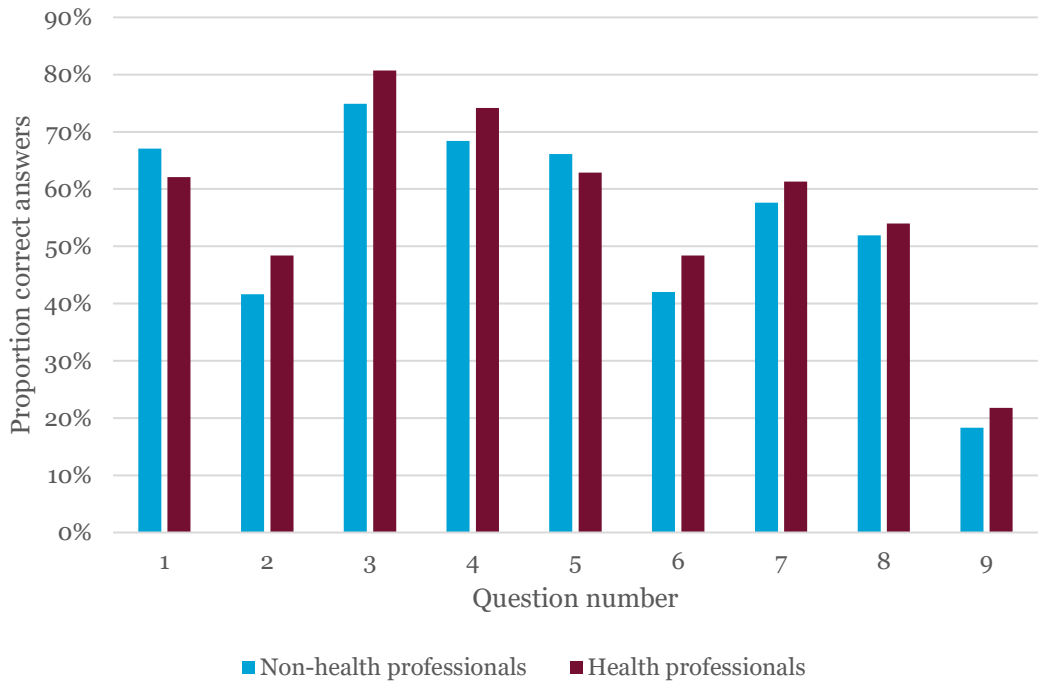
Question *	Key Concept †	Missing (n=626)	Correct (n=626)
1	3.5 Decisions about treatments should not be based on consideration only their benefits	6 (1.0%)	413 (66.0%)
2	Numeracy* and 1.3 An 'outcome' may be associated with a treatment, but not caused by the treatment	6 (1.0%)	269 (43.0%)
3	Numeracy*	5 (0.8%)	476 (76.5%)
4	2.12 Relative effects of treatments alone can be misleading	2 (0.3%)	433 (69.2%)
5	Numeracy and 2.12 Relative effects of treatments alone can be misleading	3 (0.5%)	408 (65.2%)
6	2.15 The use of p-values to indicate the probability of something having occurred by chance may be misleading; confidence intervals are more informative	9 (1.4%)	269 (43.0%)
7	2.15 The use of p-values to indicate the probability of something having occurred by chance may be misleading; confidence intervals are more informative	4 (0.5%)	365 (59.3%)
8	2.12 Relative effects of treatments alone can be misleading	1 (0.2%)	326 (52.1%)
9	1.3 An 'outcome' may be associated with a treatment, but not caused by the treatment	2 (0.3%)	120 (19.2%)



* CURE question (see Box 1)

† From a list of 34 Key Concepts that are important for people to understand and apply when assessing claims about treatment effects.[30] Numbering corresponds to numbering in the list of Key Concepts.

Figure 2. Proportions of correct answers for health professionals and non-health professionals



The results of the stepwise regression analysis are shown in Table 2. The final model, which included five variables, resulted in an adjusted R squared of 0.28. The five variables were for education level (3 variables), age, and Internet access.

Table 2. Regression analysis for total scores (n=626)

Variable	Beta	Standard error	T-statistic	P-value
1-3 years of college or university education	1.344	0.179	7.530	<0.00001
Masters level education	1.990	0.180	11.068	<0.00001
PhD education	3.359	0.591	5.682	<0.00001
Age	-0.016	0.005	-3.455	0.00059
Internet access	0.545	0.223	2.449	0.0146
Constant	4.375	0.354	12.344	

Discussion

The ability of patients to participate in decisions about treatments depends not only on the extent to which they have access to reliable information, but also on the extent to which patients can assess the trustworthiness of information about treatment effects.[31] The CURE questionnaire was developed to measure patients' ability to do this.

In our survey of Norwegian adults, the nine questions in the CURE questionnaire had low levels of missing data that all were under two percent. The total scores were reasonably normally distributed with a mean score close to answering half of the questions correctly. The comparisons with sociodemographic variables provided evidence for the validity of the questionnaire. Of the 70 comparisons, 57 (81.4%) met the a priori hypotheses for the differences between two groups or consistent trend across more than two groups; 35 (50%) were unlikely to have occurred by chance ($P<0.05$). As hypothesised, the level of education had the strongest association with scores, as indicated by the three variables representing higher levels of education in the regression equation (Table 2). All the correlations between educational level and the proportion of correct responses to individual questions were unlikely to have occurred by chance ($P<0.01$), except for question 7, which assessed understanding of confidence intervals (Table 1).

The proportions of correct answers for each question and the total scores were similar for health professionals and those who were not health professionals (Figure 2). This suggests that there are important gaps in health professionals' as well as patients' understanding of key concepts that they should understand and apply when assessing claims about treatment effects. On the other hand, younger respondents had a higher proportion of correct responses and higher total scores, suggesting that there may have been improvements over time in the ability of both health professionals and patients to assess treatment claims.

This was our first attempt in Norway to measure peoples understanding of key concepts that people need to understand to be able to assess claims about treatment effects. However, it only measured understanding of four of 34 Key Concepts that are necessary to understand and apply when assessing claims about treatment effects,[30] and three questions assessed numeracy, at least in part (Figure 1). Another disadvantage of the CURE questionnaire was that it was a fixed questionnaire. In contrast, the Claim Evaluation Tools database includes a battery of multiple choices questions covering 34 Key Concepts. These questions can be used for children or adults in tests in schools and other teaching settings and in self-assessment tools, in randomised trials evaluating the effects

of educational interventions, and in cross-sectional studies to gauge the ability of a population to assess treatment claims and make informed health choices. The questions from the CURE questionnaire addressing Key Concepts have been added to the Claim Evaluation Tools database and can be combined with other questions from this database to develop tests for measuring such ability in different populations. However, for questions taken from the Claim Evaluation Tools database to be used in formal evaluations, we recommend that these undergo Rasch analysis to ascertain the reliability and validity of those items in concordance with modern standards for such evaluations.[32,33]

The poor response rate of 25% is a limitation of this study. It is possible that those that did not respond had a lower education level than those who did. This would mean that the percentage of correct responses would have been lower if we had a more representative sample.

Conclusion

We developed the CURE questionnaire based on a review of already existing instruments, consultation with evidence-based healthcare experts, and with patients. The questionnaire includes nine questions assessing understanding of four key concepts relevant to assessing claims about treatment effects and numeracy. The findings of a survey of Norwegian adults using this questionnaire suggests that there are important gaps in both health professionals' and patients' ability to assess claims about treatment effects and make informed health choices.

References

1. Entwistle VA, Sheldon TA, Sowden A, et al. Evidence-informed patient choice. Practical issues of involving patients in decisions about health care technologies. Int J Technol Assess Health Care 1998; 14:212-25.
2. Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn't. BMJ 1996; 312:71-2.
3. Coulter A. Partnerships with patients: the pros and cons of shared clinical decision-making. J Health Serv Res Policy 1997; 2:112-21.
4. Glenton C, Paulsen EJ, Oxman AD. Portals to Wonderland: health portals lead to confusing information about the effects of health care. BMC Med Inform Decis Mak 2005; 5:7.
5. Guyatt GH, Ray J, Gibson N, et al. A journalist's guide for health stories. American Medical Writers Association Journal 1999; 14:32-41.
6. Schwartz LM, Woloshin S. The media matter: a call for straightforward medical reporting. Ann Intern Med 2002; 140:226-8.
7. Woloshin S, Schwartz LM. Press releases: translating research into news. JAMA 2002; 287:2856-8.
8. Stamm K, Williams JW Jr, Noel PH, et al. Helping journalists get it right: a physician's guide to improving health care reporting. J Gen Intern Med 2003; 18:138-45.
9. Moynihan R, Soumerai S, Fletcher R, et al. Tipsheet for reporting on drugs, devices and medical technologies. New York: Common Wealth Fund, 2004.
10. Dickersin K, Braun L, Mead M, et al. Development and implementation of a science training course for breast cancer activists: Project LEAD (leadership, education and advocacy development). Health Expect 2001; 4:213-20
11. Milne R, Oliver S. Evidence-based consumer health information: developing teaching in critical appraisal skills. Int J Qual Health Care 1996; 8:439-45.
12. Castle JC, Chalmers I, Atkinson P, et al. Establishing a library of resources to help people understand Key Concepts in assessing treatment claims - The Critical thinking and Appraisal Resource Library (CARL). PLoS One, in press.
13. Austvoll-Dahlgren A, Nsangi A, Semakula D. Interventions and assessment tools addressing key concepts people need to know to appraise claims about treatment effects: a systematic mapping review. Syst Rev 2016; 5:215.

14. Austvoll-Dahlgren A, Semakula D, Nsangi A, et al. The development of the "Claim Evaluation Tools": assessing critical thinking about effects. BMJ Open, in press.

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

16. Voss M. Checking the pulse: Midwestern reporters' opinions on their ability to report health care news. Am J Public Health 2002; 92:1158-60.

17. MacRae HM, Regehr G, Brenneman F, et al. Assessment of critical appraisal skills. Am J Surg 2004; 187:120-3.

18. Coomarasamy A, Taylor R, Khan KS. A systematic review of postgraduate teaching in evidence-based medicine and critical appraisal. Med Teach 2003; 25:77-81.

19. Godwin M, Seguin. Critical appraisal skills of family physicians in Ontario, Canada. BMC Med Educ 2003; 3:10.

20. Johnston JM, Leung GM, Fielding R. The development and validation of a knowledge, attitude and behaviour questionnaire to assess undergraduate evidence-based practice teaching and learning. Med Educ 2003; 37:992-1000.

21. Ramos KD, Schafer S, Tracz SM. Validation of the Fresno test of competence in evidence based medicine. BMJ 2003; 326:319-21.

22. Beasley BW, Woolley DC. Evidence-based medicine knowledge, attitudes, and skills of community faculty. J Gen Intern Med 2002; 17:632-9.

23. Fritsche I, Greenhalgh T, Falck-Ytter Y, et al. Do short courses in evidence based medicine improve knowledge and skills? Validation of Berlin questionnaire and before and after study of courses in evidence-based medicine. BMJ 2002; 325:1338-41.

24. Taylor R, Reeves B, Mears R, et al. Development of a questionnaire to evaluate the effectiveness of evidence-based practice teaching. Med Educ 2001; 35:544-7.

25. Baker DW, Williams MV, Parker RM, et al. Development of a brief test to measure functional health literacy. Patient Educ Couns 1999; 38:33-42.

26. Thomas KG, Thomas MR, Dupras DM. Assessment tools for evaluating critical appraisal skills. Med Educ 2004; 38:545-76.

27. Smith CA, Ganschow PS, Reilly BM, et al. Teaching evidence-based medicine skills: a controlled trial of effectiveness and assessment of durability. J Gen Intern Med 2000; 15:710-15.

28. Feinstein AR. Clinimetrics. New Haven, CT: Yale University Press; 1987; 141-66.

29. Oxman AD, Guyatt GH, Cook DJ, Jaeschke R, Heddle N, Keller J. An index of scientific quality for health reports in the lay press. J Clin Epidemiol 1993; 46:987-1001.

30. Austvoll-Dahlgren A, Oxman AD, Chalmers I, et al. Key concepts that people need to understand to assess claims about treatment effects. J Evid Based Med 2015; 8:112-25.

31. Irwig L, Irwig J, Trevena L, et al. Smart Health Choices: making sense of health advice. London: Hammersmith Press, 2008.

32. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010; 19:539-49.

33. Austvoll-Dahlgren A, Oxman AD, Chalmers I, et al. Manual for preparing a test or questionnaire based on the Claim Evaluation Tools database. 22 November 2016.

Appendix 1. Critical Use of Research Evidence (CURE)

A questionnaire to assess consumers' ability to understand and use reports about the effects of health care

Critical Use of Research Evidence (CURE) – Part I

Below are some statements. Please read each one carefully and give the answer that comes closest to the way you feel.

	(Please cross one box on each line)	I strongly disagree	I disagree	I don't agree or disagree	I agree	I strongly agree
1.	When getting advice from my doctor I would like to know about the research evidence	☐	☐	☐	☐	☐
2.	Having to make important decisions about my own health makes me uneasy	☐	☐	☐	☐	☐
3.	I do not question my doctor's judgement	☐	☐	☐	☐	☐
4.	If I get sick, health research might help me to decide what to do	☐	☐	☐	☐	☐
5.	I prefer that my doctor tells me if the benefits of a treatment are uncertain	☐	☐	☐	☐	☐
6.	When it comes to my health, I can only do what my doctor tells me to do	☐	☐	☐	☐	☐

Please try to remember the last time you searched for information, or considered a specific story about health or health care in the mass media. Then, please read each statement below carefully and tick the box that is closest to what you did.

	(Please tick one box on each line)	No	Yes
7.	I considered whether the information was based on research evidence	☐	☐
8.	I considered the quality of the evidence that the information was based on	☐	☐
9.	I considered whether the information was based on a randomised trial or systematic review	☐	☐

We will now ask you some questions about research methods and randomised controlled trials. In a randomised controlled trial patients were randomised to receive a new diet-pill or an old diet-pill for treating obesity. For each question below please tick one box to show whether you disagree or agree with the statement. Please do not guess. If you are unsure, please tick "don't know/unsure".

	(Please tick one box on each line)	**Disagree**	**Agree**	**Don't know / unsure**
10.	By randomising patients in the trial, we can be more confident that the groups are comparable	☐	☐	☐
11.	Whenever possible, participants and researchers in a study like this, should not know who receives the new treatment	☐	☐	☐
12.	Randomised trials are always high quality research	☐	☐	☐

We will now ask you some questions about research methods and systematic reviews. A systematic review was done to sum the results of available research on treatment for obesity. For each question below please tick one box to show whether you disagree or agree with the statement. Please do not guess. If you are unsure, please tick "don't know/unsure".

	(Please tick one box on each line)	**Disagree**	**Agree**	**Don't know / unsure**
13.	Systematic reviews are more useful for understanding whether a treatment works than individual studies	☐	☐	☐
14.	Information based on a systematic review is more likely to be correct than information written by an expert	☐	☐	☐
15.	Systematic reviews are always high quality research	☐	☐	☐

Below are some statements. Please read each one carefully and give the answer that comes closest to the way you feel.

	(Please cross one box on each line)	I strongly disagree	I disagree	I don't agree or disagree	I agree	I strongly agree
16.	My doctor is up to date with research about the best treatments	☐	☐	☐	☐	☐
17.	Most drugs have side effects. Decisions about whether the benefits outweigh the side effects are best left to the experts	☐	☐	☐	☐	☐
18.	Interpreting research is best left to the experts	☐	☐	☐	☐	☐
19.	Newspaper reports are a reliable source of information about health care	☐	☐	☐	☐	☐
20.	Whenever possible new treatments should be tested using randomised controlled trials	☐	☐	☐	☐	☐
21.	I am in control of decisions about my health	☐	☐	☐	☐	☐
22.	If I had to make a decision about my own health or health care I would try to use the research evidence	☐	☐	☐	☐	☐
23.	The benefits of most drugs are uncertain	☐	☐	☐	☐	☐

24. How difficult or easy do you find health care research to understand?
(Please tick one box)

- Very difficult ☐
- Fairly difficult ☐
- A little difficult ☐
- Neither difficult or easy ☐
- A little easy ☐
- Fairly easy ☐
- Very easy ☐
- Don't know ☐

Critical Use of Research Evidence (CURE) – Part II - statistics

1. You have been told by your doctor that you have a serious heart disease. Which is the most important factor when deciding to use a drug to treat the disease?
(Please tick one box)

- ☐ The number of deaths it prevents
- ☐ That the benefits outweigh the side-effects
- ☐ The number of people who have the disease
- ☐ Don't know

2. A new study finds that there is 30 deaths among people who eat broccoli regularly compared to100 deaths among people who don't eat broccoli at all.

According to this study, which statement best describes how eating broccoli relates to death?

- ☐ Lowers the risk of death
- ☐ Doesn't change the risk of death
- ☐ Raises the risk of death
- ☐ Can't tell from this information

3. About 6,500 Norwegians will be diagnosed with diabetes this year.

What is your best guess about a Norwegian's chance of being diagnosed with diabetes in the next year? *(Please tick one box)*

- ☐ 6,500 to one
- ☐ 6,500 divided by the number of Norwegians
- ☐ Don't know

4. Mrs. Jones is told by her doctor that she has a 28 in 1000 chance of dying from cancer and a 59 in 1000 chance of dying from a stroke.

Mrs Jones' doctor now tells her that a new pill, "ANTI-STROKE", will lower her chance of dying from stroke by 50%. Another pill, "ANTI-CANCER", will lower her chance of dying from cancer by 50%.

(a) She can only take 1 pill. Assuming the two pills are equally safe and the cost is the same, which do you suggest she take if her only goal is to lower her chance of dying? *(Please tick one box)*

ANTI- STROKE pill ☐

ANTI- CANCER pill ☐

Either pill, it makes no difference ☐

Don't know ☐

(b) Mrs Jones decides to take the ANTI-CANCER pill. Now, what is her chance of dying from cancer? *(Please tick one box)*

0 in 1000 ☐

7 in 1000 ☐

14 in 1000 ☐

21 in 1000 ☐

Don't know ☐

5. What do we mean when we say that the result of the study of a treatment is "statistically significant"? *(Please tick one box)*

That the result is important ☐

That the results are unlikely to be due to chance ☐

That the results represent a big difference ☐

Don't know ☐

6. What does a confidence interval tell us about a treatment effect?
(Please tick one box)

- It is the area where the result is unlikely to be due to systematic error ☐
- It is the area where the treatment effect is so big that it should be used in clinical practice ☐
- It is the area where it is most probable that we will find the true treatment effect ☐
- Don't know ☐

7. Which of the following programmes for people who have had a heart attack is most effective? *(Please tick one box)*

In the three years after a heart attack:

- Programme A would reduce the rate of deaths by 20% ☐
- Programme B would reduce the number of deaths by 3% ☐
- Programme C would increase patient survival from 84% to 87% ☐
- Programme D would mean that 31 people needed to enter the programme to prevent one death ☐
- They are all equally effective ☐
- Don't know ☐

8. A new study found that people drinking alcohol regularly were more likely to have cancer.

According to this study, which statement best describes how drinking alcohol relates to cancer? *(Please tick one box)*

- Alcohol lowers the risk of cancer ☐
- Alcohol doesn't change the risk of cancer ☐
- Alcohol raises the risk of cancer ☐
- Can't tell from this information ☐

Below are some statements. Please read each one carefully and give the answer that comes closest to the way you feel.

	(Please tick one box on each line)	**I strongly disagree**	**I disagree**	**I don't agree or disagree**	**I agree**	**I strongly agree**
9.	I understand the numbers reported in health research results in the media	☐	☐	☐	☐	☐
10.	To make a good decision about my health, it is important to know how to interpret the numbers reported in health research results in the media	☐	☐	☐	☐	☐
11.	I want my doctor to give me numbers when explaining treatments to me	☐	☐	☐	☐	☐

12. PERSONAL CHARACTERISTICS

Are you female or male?

☐ Female ☐ Male

Please write you age below:

☐☐ years

What is your level of education (please tick one box)?

☐ High school/ elementary school

☐ 1- 3 years of College or university

☐ More than 3 years of college or university

☐ PhD

Do you work within the health field?

☐ No ☐ Yes

Do you have any research experience?

☐ No ☐ Yes

Do you have Internet access?

☐ No ☐ Yes

Thank you for participating!