© Health Research and Educational Trust
DOI: 10.1111/1475-6773.12311
RESEARCH BRIEF

The Role of Outcome Forecasts in Patients' Treatment Decisions—Evidence from a Survey Experiment on Knee Replacement Surgery

Iris Kesternich, Francis G. Caro, Alison S. Gottlieb, Susanne Hoffmann, and Joachim K. Winter

Objective. To analyze the effect of personalized outcome probabilities on treatment decisions.

Data Source. American Life Panel.

Study Design. We designed an experimental survey that respondents answered by recommending replacement surgery for hypothetical patients diagnosed with knee osteoarthritis. The primary manipulation was whether respondents were given personalized outcome probabilities for the patients.

Principal Findings. Patient-specific outcome forecasts significantly influenced decisions, with effect sizes comparable to those of physicians' opinions and patient testimonials. The presentation format also affected the probabilities' influence on choices. We did not find evidence that the educational level of subjects affected the weight that they give to outcome probabilities in making recommendations.

Conclusion. Our study points to a need for more research on the use of patient-specific outcome forecasts.

Key Words. Patient treatment choice, decision aids, outcome probabilities, hypothetical choice experiment

To strengthen patient decision making, decision aids have been introduced for many medical procedures. They are designed to complement physician recommendations. Decision aids inform patients about the “health condition, the [treatment] options, associated benefits, harms, probabilities, and scientific uncertainties” (O'Connor et al. 2009). Systematic reviews of decision aids show that they frequently include statistical outcome probabilities, for example, in 80 percent of the studies reviewed in O'Connor et al. (2009) and in 86 percent of the studies reviewed in Stacey et al. (2011).

However, the outcome probabilities typically included in decision aids are constant across patients. As individual patient characteristics can influence outcome probabilities, personalized outcome probabilities might have stronger impacts on patients' decision making than general outcome probabilities. Providing patients with personalized probabilities is a realistic future possibility because of advances in information technology and the potential availability of medical outcome data aggregated by insurance providers (see, e.g., Heiss et al. 2013).

In this paper, we analyzed the relative importance of person-specific probabilities relative to other sources of information. We conducted a hypothetical survey experiment to measure the relationship between personalized outcome probabilities and treatment choices for knee osteoarthritis.

Knee osteoarthritis is a chronic illness well suited for studying hypothetical treatment decisions: it is a common condition with a strong negative effect on mobility and the quality of life of older people. Around 30 percent of the elder population suffer from this disease (Felson et al. 1987), and over 9.3 million adults had symptoms of knee osteoarthritis in 2005 (American Academy of Orthopedic Surgeons [AAOS] 2009).

Several forms of treatment are available that vary widely in the commitment required of patients and their likely therapeutic value. Total knee replacement can improve the patient's condition permanently but is the most expensive option, carries the risk of infections, and involves lengthy, painful rehabilitation. Therefore, knee replacements should be considered carefully not only because of its rapidly increasing aggregated costs (Tomek et al. 2012), but also for its implications for patients' well-being.

The success of a total knee replacement depends strongly on patient attributes—on medical factors such as the degree of pain or patient weight as well as on sociodemographic variables such as age and the support by one's social environment (Hawker et al. 2000, 2002; Kennedy et al. 2003; Escobar et al. 2007; Gandhi et al. 2010). Thus, decision aids might be improved by including outcome probabilities that are conditioned on the patient attributes (and not just population averages).

Address correspondence to Iris Kesternich, Dr. oec. publ., Department of Economics, University of Leuven, Naamsestraat 69, 3000 Leuven, Belgium; e-mail: iris.kesternich@kuleuven.be. Francis G. Caro, Ph.D., and Alison S. Gottlieb, Ph.D., are with the Gerontology Institute, University of Massachusetts Boston, Boston, MA. Susanne Hoffmann, Dr. oec. publ., and Joachim K. Winter, Dr. rer. pol., are with the Department of Economics, University of Munich, Munich, Germany.

Respondents were asked to recommend knee replacement surgery for hypothetical patients. The primary manipulation was whether respondents were given personalized statistical outcome probabilities for the hypothetical patients. Survey data were collected using a large, nationally representative Internet survey.

METHODS

Sample

The respondents to our Internet survey were members of the RAND American Life Panel (ALP, mmicdata.rand.org/alp/) aged 50 and older. The ALP panel seeks to be representative of the adult U.S. population (<https://mmicdata.rand.org/alp/index.php?page=panel> or Heiss et al. 2013, for a comparison with the Health and Retirement Survey). Members of the ALP are surveyed periodically in an Internet format and receive modest financial compensation for participation. We had access to respondents' standard demographic data obtained from previous surveys. The ALP invited all 2296 members aged 50 and older who had not participated in our pilot study to participate in our survey. A total of 1616 members of the ALP participated in our study. Respondents were fairly evenly represented by men and women and reflected a broad age range (between 50 to 93 years) where knee osteoarthritis is a common health condition. More information on our sample can be found in Hoffman et al. (2014).

Study Design

We conducted a stated-choice survey (Louviere, Hensher, and Swait 2000) and used a fractional factorial design based on Rossi and Anderson (1982) to construct hypothetical choice situations.

Subjects were randomly assigned to groups that differed in whether and how information on personalized outcome probabilities was provided before a treatment decision. Use of the Internet to administer the study allowed us to engage respondent interest by using videos to show hypothetical patients and physicians.

The participants were asked to provide a recommendation for or against total knee replacement surgery for a hypothetical patient with randomly varying attributes.

In total, each respondent was asked to provide recommendations for three vignettes differing in several dimensions. Four of these dimensions were central to our analysis of decision making processes of patients:

1. Person-specific outcome prediction delivered by an online tool (optional; levels: above-average or below-average chances for successful surgery)
2. Physician recommendations (mandatory; levels: positive or neutral)
3. Second opinion by a specialist (optional; levels: strong recommendation, substantial reservations)
4. Patient testimonials (one testimonial mandatory, one testimonial optional; levels: positive or negative; shown only with the third vignette)

For outcome predictions, second opinions, and patient testimonials, respondents had a choice of whether they wanted to be provided with the respective information or not. Thus, we explicitly recognized that patients can be selective in the information they consider when using decision aids. In addition, the chosen options enabled us to spotlight respondent interest in various information dimensions. In addition, we asked respondents how interested they were in the different information dimensions after the experiment. They were also asked to rate the relative importance of various vignette dimensions in choice making to check our results methodologically.

Outcome probabilities were presented in one of the three randomly assigned formats: verbal (higher/lower than the average patient), numeric (actual percentages were provided for both the average and own outcome probabilities), or numeric accompanied by a bar graph.

The experiment proceeded as follows: through a video-clip, respondents experienced a mini-lecture on knee osteoarthritis by a physician, with surgery included among the treatment options. The lecture's content was drawn largely from a booklet and DVD on knee osteoarthritis published by Health Dialog (2012).

Respondents were then asked to review the situation of one vignette person who was considering knee replacement surgery to treat knee osteoarthritis. All respondents were shown one surgeon's recommendation and one patient testimonial. Respondents could choose to see a second patient testimonial, a second physician's opinion, and one outcome prediction delivered by an online tool.

RESULTS

Respondents tended to request all of the information offered to assist them in their decision making. Respondents most often requested outcome probabilities (90 percent) and second opinions (88 percent); 74 percent requested patient testimonials.

To assess the relative importance of the various information sources (surgeon's recommendation, specialist's second opinion, patient testimonials, and personalized outcome probabilities) on the probability of recommending surgery, we ran a linear regression with a binary variable indicating whether surgery was recommended in different scenarios using the 1614 respondent-vignette observations for the third vignette (the only vignette that included patient testimonials). In the regression, the predictors were indicators for the various information sources as well as demographic and health controls. The results of this linear probability model are reported in Table 1 (see below). We used a linear probability model for ease of interpretation, but results were qualitatively similar when using a probit or logit model.

Outcome probabilities significantly influenced surgery recommendations. An above-average outcome probability increased the probability of recommending surgery by 16 percentage points; a below-average outcome probability led to a decrease of 12 percentage points. The magnitude of the influence of outcome probabilities was similar to that of a physician's recommendation (an 18 percentage point increase in case of a positive recommendation), a specialist's second opinion (an 11 [22] percentage points increase [decrease] in case of a positive [negative] opinion), and patient testimonials (a 15 [13] percentage points increase [decrease] for two positive [negative] testimonials).

As a methodological check, we compared the experimental results to respondents' answers to the questions asked at the end of the survey regarding which information dimensions were most influential for their decisions. Respondents most often reported outcome probabilities as most influential (27 percent). The physician's recommendation and the second opinion were reported as most influential by 13 percent and 14 percent of respondents, respectively.

For each dimension, respondents were also asked how interested they were in considering this information (on a scale from "very interested" to "not at all interested"). Here, 67 and 64 percent of respondents stated that they were "very interested" in outcome probabilities and second opinions, respectively.

Table 1: Linear Regressions Predicting Surgery Recommendation

<i>Dependent Variable: Variables</i>	<i>Surgery Recommendation</i>		
	(1)	(2)	(3)
Positive surgeon's recommendation	0.177*** [0.025]	0.175*** [0.025]	0.177*** [0.025]
Second opinion			
Strong recommendation	0.106*** [0.030]	0.112*** [0.030]	0.110*** [0.030]
Neutral view	-0.215*** [0.029]	-0.212*** [0.029]	-0.214*** [0.030]
Second opinion not chosen	0.050 [0.045]	0.052 [0.046]	0.052 [0.045]
Person-specific outcome forecast			
Forecast above average	0.159*** [0.028]	0.086** [0.041]	0.095** [0.039]
Forecast above average × Numeric format		0.088* [0.047]	
Forecast above average × Graphic format		0.120*** [0.046]	
Forecast below average	-0.115*** [0.030]	-0.149*** [0.040]	-0.178*** [0.039]
Forecast below average × Numeric format		0.068 [0.050]	
Forecast below average × Graphic format		0.036 [0.047]	
Forecast not chosen	-0.025 [0.049]	-0.025 [0.050]	0.000 [0.058]
Numeric forecast format			0.077 [0.034]
Graphic forecast format			0.079** [0.033]
Forecast above average × Low education			0.047 [0.066]
Forecast below average × Low education			0.048 [0.067]
Forecast not chosen × Low education			-0.061 [0.103]
Patient testimonials			
Two positive testimonials	0.148*** [0.044]	0.142*** [0.044]	0.142*** [0.044]
Two negative testimonials	-0.125*** [0.041]	-0.131*** [0.041]	-0.130*** [0.041]
Mixed message testimonial	0.031 [0.028]	0.028 [0.028]	0.027 [0.028]
One positive testimonial	0.068* [0.038]	0.069* [0.038]	0.067* [0.038]
One negative testimonial	-0.053 [0.045]	-0.061 [0.045]	-0.063 [0.045]
Respondent's background characteristics			
Gender	0.042* [0.023]	0.039* [0.023]	0.042* [0.023]
High school or less	0.051* [0.027]	0.051* [0.027]	0.024 [0.051]

Notes: Coefficients obtained by Ordinary Least-Squares estimation; standard errors in brackets
 *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The testimonials elicited significantly less interest from respondents (44 percent were “very interested” in them).

Finally, respondents were asked to choose whether second opinions or outcome probabilities were more helpful for making recommendations. Exactly, 58 percent selected outcome probabilities, while 42 percent chose second opinions.

Other investigators have raised concerns that outcome probabilities might be less helpful for patients with low education or low numeracy skills (Peters et al. 2007). Thus, we stratified our sample into those with high school degrees or less (24 percent of our respondents) and those with higher educational degrees. Of respondents in the lower education category, 14.9 percent chose not to view the probabilities, while only 8.6 percent of the higher education category chose not to view them ($p < .001$). However, respondents from the lower education category showed no significant differences in the way that the probability influenced their treatment choice. An F-test for the joint significance of the education interaction in specification 3 has a p -value of .250 (see specification 3 in Table 1).

Because presentation style can matter greatly for how respondents understand and interpret probabilities, we also analyzed the influence of presenting probabilities as text, graphically, or numerically. We found that both the graphic and numeric formats for presenting above-average outcome probabilities led to a higher probability of recommending surgery than text versions. An F-test for the joint significance of the graphic and numeric interactions has a p -value of .065.

DISCUSSION

Survey respondents gave knee replacement recommendations for hypothetical patients suffering from knee osteoarthritis; information on these patients was manipulated experimentally. Our findings showed that there is great interest among patients in using outcome probabilities to assist them in making decisions.

Designers of web-based decision aids may increasingly be able to offer personalized outcome probabilities. Results of treatment outcome studies for various chronic illnesses have become available from large representative patient samples. Moreover, for some common health procedures, there is a sufficient body of outcome studies so that multivariate statistical techniques can be used to predict how particular configurations of independent variables

affect outcome probabilities. For example, the Informed Medical Decisions Foundation provides a cardiac risk calculator on its homepage (www.cardiac-riskcalculator.org). Fagerlin et al. (2010, 2011) used an online tool that provided personalized predictions for 5-year risk of breast cancer. These outcome studies made it feasible to offer reliable outcome forecasts that control for a number of important variables. As a result, decision-aid websites that feature interactive user interfaces permit patients to provide information about themselves that can serve as the basis for patient-specific outcome forecasts.

A number of questions about the use of personalized outcome forecasts in decision aids deserve attention. Among them is what covariates (such as age, gender, race, socioeconomic status, severity, and comorbidities) should be used routinely in estimating outcomes. This question can best be answered by statistical analysis of large-scale datasets. Another question meriting attention is whether patients should be informed about the margin of error in the outcome forecasts.

In our discussion, we have concentrated on patient-specific probabilities, but physician-specific outcome probabilities could be valuable as well. A summary of the literature on quality reporting in health economics can be found in Dranove (2012). The evidence on quality reporting suggests that it does not only help patient choice but also increases average quality in health care markets.

However, development of sound patient-specific outcome models faces several technical obstacles. Until now, it has been impossible for the Informed Medical Decisions Foundation to provide personalized prognoses for diabetes because there is no general agreement on which risk factors should be included and how they should be weighted. For knee replacement surgery due to osteoarthritis, for example, there is no national effort to track outcomes after knee replacement procedures in the United States, and only few population-level outcome studies exist.

Also, population data often lack coverage of certain subgroups of patients. Those patients who face most medical decisions, elders, are often not included in medical trials; the same holds for children and patients with comorbidities (Tomek et al. 2012). However, these limitations also affect aggregate outcome probabilities: if aggregate probabilities are calculated based on studies on the nonelder population only, then the aggregate success probability may be overestimated for an older person facing a treatment decision. Further interest in using outcome probabilities to supporting patient decision making is likely to stimulate efforts to obtain strong outcome data on major discretionary treatment strategies. Major insurers have reason to support outcome

research that would provide the basis for high quality personalized outcome forecasts in decision aids.

There are concerns that patients, especially those with low education, might have difficulties in interpreting numerical information and probabilities. This concern about implications of education applies to both population outcome forecasts and personalized ones. Personalized probabilities might enable patients with low levels of education to access information more easily as using aggregate data requires the cognitive ability to apply general probabilities to one's own circumstances. It is assuring that there were no significant differences in how outcome probabilities influenced the decisions of respondents from the high and low education categories. Our results suggest that less educated people might need stronger incentives to access and use probabilities (even though over 85 percent of the lower education category chose to consider the outcome forecasts). However, it might be difficult to generalize our findings as all our respondents were able to finish an Internet-administered survey.

Furthermore, probabilities communicated in both graphic and numeric formats were found to have more influence than probabilities presented solely in a text format. This indicates the importance of presenting probabilities in a way that makes them accessible to less well-educated people (Peters et al. 2007).

In conclusion, there is a strong case for conducting more research on the contribution of patient-specific outcome forecasts to decision aids. The new Patient-Centered Outcomes Research Institute, authorized by Congress through the Patient Protection and Affordable Care Act of 2010, might support the outcome research needed for sound personalized outcome forecasts in decision aids.

Limitations

One limitation of our study is that respondents made hypothetical, not actual, choices. In real choices, physicians' opinions might carry more weight, and it might be easier to make a hypothetical decision to go through surgery than to go through with it in real life. However, a random assignment of information to patients is impossible to implement in real patient decisions for ethical reasons.

We were concerned entirely with patient-specific outcome forecasts. Consequently, we cannot estimate differences in the impact of personalized and general outcome forecasts. Further research examining the differential impact of personalized and general outcome forecasts would be useful.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: The research was supported in part by Investigator initiated grant 0212-1 from the Informed Medical Decisions Foundation. Maria E. Suarez-Almazor, M.D., Ph.D., served as medical consultant for the project. Arie Kapteyn, Ph.D., Tania Gutsche, and other staff members of the RAND Corporation's American Life Panel arranged for production of the video files, programmed and administered the experiment, and prepared the data files. Felix Ferstl provided valuable research assistance.

Disclosures: None.

Disclaimers: None.

REFERENCES

- American Academy of Orthopedic Surgeons [AAOS]. 2009. "Arthritis and Related Conditions." AAOS Now March 2009 Issue [accessed on June 12, 2013]. Available at <http://www.aaos.org/news/aaosnow/mar09/research6.asp>
- Dranove, D. 2012. "Health Care Markets, Regulators, and Certifiers." In *Handbook of Health Economics*, edited by M. V. Pauly, T. G. McGuire, and P. P. Barros, pp. 639–690. Amsterdam, The Netherlands: North-Holland.
- Escobar, A., J. M. Quintana, A. Bilbao, J. Azkárate, J. I. Güenaga, J. C. Arenaza, and L. F. Gutierrez. 2007. "Effect of Patient Characteristics on Reported Outcomes after Total Knee Replacement." *Rheumatology* 46 (1): 112–9.
- Fagerlin, A., A. J. Dillard, D. M. Smith, B. J. Zikmund-Fisher, R. Pitsch, J. B. McClure, S. Greene, S. Hensley Alford, V. Nair, D. F. Hayes, C. Wiese, and P. A. Ubel. 2011. "Women's Interest in Taking Tamoxifen and Raloxifene for Breast Cancer Prevention: Response to a Tailored Decision Aid." *Breast Cancer Research and Treatment* 127 (3): 681–8.
- Fagerlin, A., B. J. Zikmund-Fisher, D. M. Smith, V. Nair, H. A. Derry, J. B. McClure, S. Greene, A. Stark, S. Hensley Alford, P. Lantz, D. F. Hayes, C. Wiese, S. Claud Zweig, R. Pitsch, A. Jankovic, and P. A. Ubel. 2010. "Women's Decisions Regarding Tamoxifen for Breast Cancer Prevention: Responses to a Tailored Decision Aid." *Breast Cancer Research and Treatment* 119 (3): 613–20.
- Felson, D. T., A. Naimarek, J. Anderson, L. Kazis, W. Castelli, and R. F. Meenan. 1987. "The Prevalence of Knee Osteoarthritis in the Elderly. The Framingham Osteoarthritis Study." *Arthritis & Rheumatism* 30 (8): 914–8.
- Gandhi, R., H. Dhotar, F. Razak, P. Tso, J. R. Davey, and N. N. Mahomed. 2010. "Predicting the Longer Term Outcomes of Total Knee Arthroplasty." *The Knee* 17 (1): 15–8.
- Hawker, G. A., J. G. Wright, P. C. Coyte, J. I. Williams, B. Harvey, R. Glazier, and E. M. Badley. 2000. "Differences between Men and Women in the Rate of Use of Hip and Knee Arthroplasty." *New England Journal of Medicine* 342 (14): 1016–22.

- Hawker, G. A., J. G. Wright, R. H. Glazier, P. C. Coyte, J. I. Williams, and E. M. Badley. 2002. "The Effect of Education and Income on Need and Willingness to Undergo Total Joint Arthroplasty." *Arthritis and Rheumatism* 46 (12): 3331–9.
- Health Dialog. 2012. "Treatment Choices for Knee Osteoarthritis" [accessed on June 10, 2013]. Available at <https://www.healthcrossroads.com/content/1a6532e8-f545-4b72-bd5a-850493b9ad1a/63e9342d-75e5-476a-b928-2e54cd141ec0/KO-A001B%20v03.pdf>
- Heiss, F., A. Leive, D. McFadden, and J. Winter. 2013. "Plan Selection in Medicare Part D: Evidence from Administrative Data." *Journal of Health Economics* 32 (6): 1325–44.
- Hoffman, S., F. G. Caro, A. S. Gottlieb, I. Kesternich, and J. K. Winter. 2014. "Contributions of Second Opinions, Outcome Forecasts, and Testimonials to Patient Decisions about Knee-Replacement Surgery." *Medical Decision Making* 34 (5): 603–14.
- Kennedy, L. G., J. H. Newman, C. E. Ackroyd, and P. A. Dieppe. 2003. "When Should We Do Knee Replacements?" *The Knee* 10 (2): 161–6.
- Louviere, J. J., D. A. Hensher, and J. D. Swait. 2000. *Stated Choice Methods: Analysis and Application*. Cambridge, UK: Cambridge University Press.
- O'Connor, A. M., C. L. Bennet, D. Stacey, M. Barry, N. F. Col, K. B. Eden, V. A. Entwistle, V. Fiset, M. Holmes-Rovner, S. Khangura, H. Llewellyn-Thomas, and D. Rovner. 2009. "Decision Aids for People Facing Health Treatment or Screening Decisions (Review)." *Cochrane Library* 3: CD001431.
- Peters, E., J. Hibbard, P. Slovic, and N. Dieckmann. 2007. "Numeracy Skill and the Communication, Comprehension, and Use of Risk-Benefit Information." *Health Affairs* 26 (3): 741–8.
- Rossi, P. H., and A. Anderson. 1982. "The Factorial Survey Approach: An Introduction." In *Measuring Social Judgements: The Factorial Survey Approach*, edited by P. H. Rossi, and S. L. Nock, pp. 16–65. Beverly Hills, CA: Sage.
- Stacey, D., C. L. Bennett, M. J. Barry, N. F. Col, K. B. Eden, M. Holmes-Rovner, H. Llewellyn-Thomas, A. Lyddiatt, F. Légaré, and R. Thomson. 2011. "Decision Aids for People Facing Health Treatment or Screening Decisions (Review)." *Cochrane Library* 3: CD001431.
- Tomek, I. M., A. L. Sabel, M. I. Froimson, G. Muschler, D. S. Jevsevar, K. M. Koenig, D. G. Lewallen, J. M. Naessens, L. A. Savitz, J. L. Westrich, W. B. Weeks, and J. N. Weinstein. 2012. "A Collaborative of Leading Health Systems Finds Wide Variations in Total Knee Replacement Delivery and Takes Steps to Improve Value." *Health Affairs* 31 (6): 1329–38.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Appendix SA2: Additional Linear Regression Results, Dependent Variable: Surgery Recommendation.

Appendix SA3: Probability Formats.