

Contributions of Second Opinions, Outcome Forecasts, and Testimonials to Patient Decisions about Knee Replacement Surgery

Susanne Hoffmann *Dr. oec. publ.*, Francis G. Caro, PhD, Alison S. Gottlieb, PhD, Iris Kesternich, *Dr. oec. publ.*, Joachim K. Winter, *Dr. rer. pol.*

Background. Decision aids are now a well-established means of supporting patients in their medical decision making. The widespread use of decision aids invites questions about how their components contribute to patient decisions. **Objective.** The objective of this study was to measure the importance of second opinions, patient-specific outcome forecasts, and patient testimonials relative to patient clinical and socioeconomic factors and the primary physician recommendation on the decision to undergo full knee replacement surgery to treat knee osteoarthritis. **Methods.** Middle-aged and older members of the RAND American Life Panel (N = 1616) chose whether to recommend surgery as a treatment for each of 3 hypothetical patients (vignettes) presented in a video-enhanced

internet survey. Vignettes randomly sampled levels of scenario attributes. **Results.** Second opinions, person-specific outcome forecasts, and 2 consistent patient testimonials strongly affected respondents' decision making; a single testimonial, however, did not significantly affect decisions. **Conclusions.** Information provided in a decision aid, including person-specific outcome forecasts and testimonials, can affect patient choices. The strong effect of testimonials and respondents' interest in reviewing them reinforces concerns about unwanted influence when testimonials are biased. **Key words:** medical decision making; knee replacement surgery; stated-choice experiment; decision guides; person-specific outcome forecasts; patient testimonials. (*Med Decis Making* 2014;34:603-614)

Decision aids are a well-established means of supporting patient decision making. According to International Patient Decision Aids Standards (IPDAS), more than 500 decision aids are available

or are being developed worldwide.¹ Two forms of information are particularly challenging for designers of decision aids: patient testimonials and outcome forecasts. Testimonials engage patient interest² but can introduce bias if patients generalize inappropriately from the small number of testimonials that can be incorporated in a decision aid.³ Personalized outcome forecasts based on large-sample research can be very informative, but patients often disregard or misunderstand probability statements used to communicate forecasts.⁴ Moreover, patients evaluate the information in decision aids along with their own experience of illness, socioeconomic factors, and interactions with their physicians.

In this study of how patients might weigh and integrate this information, we focus on patients considering full knee replacement surgery to treat knee osteoarthritis. Osteoarthritis of the knee is a widespread chronic condition among middle-aged and older people that has several treatment options, including surgery. Full knee replacement surgery is often highly effective, but the procedure is expensive and requires active patient cooperation during an extended and often painful rehabilitation. For many

Received 17 December 2012 from University of Munich, Munich, Germany (SH, IK, JKW), and University of Massachusetts Boston, Boston, MA (FGC, ASG). Maria E. Suarez-Almazor, MD, PhD served as medical consultant for the project. Arie Kapteyn, PhD, Tania Gutsche, and other staff 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. Financial support by the Informed Medical Decisions Foundation, Boston, MA, is gratefully acknowledged. Revision accepted for publication 17 February 2014.

Supplementary material for this article is available on the *Medical Decision Making* Web site at <http://mdm.sagepub.com/supplemental>.

Address correspondence to Francis G. Caro, Gerontology Institute, University of Massachusetts Boston, 100 Morrissey Blvd, Boston, MA 02125; e-mail: frank.caro@umb.edu.

© The Author(s) 2014

Reprints and permission:

<http://www.sagepub.com/journalsPermissions.nav>

DOI: 10.1177/0272989X14527796

patients, nonsurgical strategies are satisfactory in controlling pain and restoring adequate physical functioning. Full knee replacement surgery is thus a good example of a discretionary medical treatment with a strong need for well-informed patient decisions.

Of special interest in this study is understanding how decisions are influenced by 3 forms of supplementary information: a specialist's second opinion, patient-specific outcome forecasts, and patient testimonials. Patient pain, opportunity costs, and the recommendation of the attending physician serve as the context for the evaluation of the impact of the supplementary information. When people are forced to make choices that involve consideration of information on multiple dimensions, the relative impact of each of the dimensions tends to be complex; further, people often have limited insight into the basis for their choices.^{5,6} We use a stated-choice experiment to vary information attributes in a controlled fashion in order to identify their independent effects on the decision.

METHODS

Study Design

The study used a random fractional factorial survey design⁷ in a stated-choice (or discrete choice) task.⁸ Each respondent was asked to provide a treatment recommendation for 3 different patient vignettes. These descriptions of hypothetical patients vary in attributes of the patient and the supplementary information provided. We randomly selected 3 vignettes from the full factorial design to present to each respondent so that each combination of attributes was equally likely to be presented.

Although this design is less efficient than partial factorial designs constructed to test particular effects,⁹ it makes no assumptions about the relationships among attributes.

We used a video-enhanced vignette design developed by the investigators.¹⁰ We delivered the information in the vignettes to respondents largely through audio and video recordings, in which a narrator read text aloud or an actor played the part of a vignette person, a doctor, or another patient providing a testimonial. The technique helped to approximate real-life situations and enabled us to offer respondents with interactive options. To control for possible mode effects, a random subsample of participants received a text-only version of our survey.

Experimental Intervention

The survey experiment proceeded as follows: First, respondents watched a video mini-lecture (about 2.5 minutes) by a physician on osteoarthritis of the knee, with surgery included among the treatment options. Content information on knee osteoarthritis was drawn largely from a booklet and DVD on knee osteoarthritis published by Health Dialog.¹¹ Respondents were then asked to review the situation of 3 randomly chosen vignette persons who were considering knee replacement surgery to treat knee osteoarthritis and who introduced themselves with their first names. After reviewing each vignette person, respondents were asked to recommend whether the vignette person should have knee replacement surgery by answering "yes" or "no" to the question "Do you recommend that [Name] have full knee replacement surgery now?"

Vignette Dimensions

Each vignette presented a different (randomly chosen) person portrayed by 1 of 6 actors (3 male, 3 female). Vignettes were presented with randomly chosen levels of up to 8 substantive dimensions.

All vignettes included information on patient pain (high or moderate), patient employment and opportunity cost associated with surgery (employed or retired crossed with low or high opportunity cost), and recommendation of the attending orthopedic surgeon (a positive surgical recommendation or a neutral recommendation). We included pain because we expected that surgery would more often be recommended by respondents when vignette persons were experiencing more intense pain. We included employment and opportunity costs because we expected that greater disruption of activities during recovery and rehabilitation would decrease receptivity to surgery. For those employed, we focused on employment income that might be lost during recovery from surgery. For retired people, we described ways in which life styles would be affected by surgery and rehabilitation. If the level of opportunity costs was high, the vignette person would experience substantial financial hardship or lifestyle disruption if surgery was chosen. We included surgeon recommendations because these often constitute the starting point for a medical decision and heavily influence patients.¹² In the neutral condition, surgery was described by the orthopedic surgeon as an acceptable option, but the emphasis was on the need for the patient to make the decision. A positive

recommendation was a statement that the surgeon recommended surgery and believed no other option would provide as much pain relief.

Four additional dimensions varied supplementary sources of information that constitute the core of our research agenda. Each of these dimensions was either presented (approximately 75% of the time) or omitted from each vignette.

Specialists second opinion. Patients are frequently encouraged to seek a second opinion from another physician before making decisions.^{13,14} Reviewing the consulting specialist's second opinion was optional and was introduced using the following statement:

[Name] has sought a second opinion from a consulting physician who has carefully reviewed [name's] health history and the proposed knee replacement operation. Do you want to find out what the specialist has to say?

The consulting specialist either strongly recommended surgery or raised substantial reservations about surgery.

Person-specific outcome forecasts. We included outcome forecasts because they are widely used in decision guides. Large-scale outcome studies for common procedures like knee replacement surgery could provide the basis for person-specific outcome projections that could control for such variables as patient age and body mass index. In our study, these person-specific outcome forecasts were presented via a tool on the homepage of a hypothetical knee osteoarthritis patient aid group and were introduced to the respondents in the following way:

[Name] also sought further information on the internet. The homepage of a nonprofit knee osteoarthritis patient aid group offers a tool that predicts likely surgery outcomes. This application uses [name's] personal and health information and compares these with a large sample of full knee replacement surgery patients to predict [name's] chances of a successful surgery outcome. The development of this tool was funded by the US Department of Health and Human Services.

Respondents were then shown a screen informing them that the vignette person's chance of a successful surgery outcome was either "above average" or "below average." The chances of success were described in 1 of 3 ways (chosen at random): only in verbal terms, in numeric terms including specific percentages, or in a graphic format that used a bar chart to represent the vignette person's chances of a successful

surgery outcome relative to the average chances of a successful surgery outcome. Outcome forecasts were available and were presented (75%) or omitted (25%) in the first and second vignettes; for the third vignette person, when forecasts were available (75%), reviewing this information was optional.

Patient testimonials. In the third vignette, we also included patient testimonials (75% of the time) because there is evidence that personal anecdotes influence decision making, testimonials are often used in decision guides, and testimonials can introduce bias when not representative of the population of patients.³ One patient testimonial was presented with the following introduction:

[Name's] friend recommended [name] view videos of real patients talking about their experiences with knee surgery. [Name] viewed the following testimonial on the internet.

Respondents who saw the first testimonial were also given the option of seeing a second testimonial. Each testimonial could be either positive or negative (chosen at random).

Therefore, in the third vignette, respondents could see no testimonials, 1 positive or negative testimonial (having chosen not to see the second testimonial), or 2 positive and 2 negative or 1 positive and 1 negative testimonial.

To minimize the potential for recency effects, respondents had an opportunity before making their recommendations to view a summary of the information on all of the dimensions except testimonials.

Design overview. Table 1 summarizes the design, listing all dimensions, their levels, and the percentages with which these levels should appear.

Appendix A illustrates the full design for 1 vignette person, including the actual wording used in the experiment for each level within the dimensions. Appendix B summarizes the proportion of the study sample who received each level of each dimension. (The online appendixes for this article are available on the *Medical Decision Making* Web site at <http://mdm.sagepub.com/supplemental>.)

Study Sample

Study participants were members of the RAND American Life Panel (ALP), a probability-based online panel that seeks to be representative of the US adult population.¹⁵ Such panels compare well with population-representative telephone surveys in generalizability and data quality.^{16,17} Members of

Table 1 Dimensions, Levels, and Level Frequency

Dimensions	Levels	Frequency (%)
Pain	High	50
	Moderate	50
Opportunity costs and employment status	High (employed)	25
	Low (employed)	25
	High (retired)	25
	Low (retired)	25
Orthopedic surgeon's recommendation	Positive	25
	Neutral	75
Specialist's second opinion	Strong recommendation	37.50
	Substantial reservation	37.50
	Not available	25
Person-specific outcome forecasts	Above average	37.50
	Below average	37.50
	Not available	25
First patient testimonial	Positive 1	18.75
	Positive 2	18.75
	Negative 1	18.75
	Negative 2	18.75
	Not available	25
Second patient testimonial	Positive 1	18.75
	Positive 2	18.75
	Negative 1	18.75
	Negative 2	18.75
	Not available	25

the ALP are drawn from the general population, are surveyed periodically online, and receive modest financial compensation for participating in particular studies.

The ALP invited all 2296 of its members aged 50 and older at the time of the survey who had not participated in our pilot study to participate in this study. A total of 1675 interviews were started; 1622 were completed. This results in a completion rate (AAPOR RR5)¹⁸ of 70.6%. * Background characteristics were incomplete for 6 respondents, who were subsequently dropped from the analysis. This resulted in a total respondent sample size of 1616.

We had access to standard demographic information about respondents obtained previously based on their participation in the panel. Additionally, we asked respondents to answer questions concerning their own medical histories with respect to chronic knee pain, knee osteoarthritis, and full knee replacement surgery as well as their friends' and close relatives' experiences with knee osteoarthritis

and full knee replacement surgery. Background characteristics of the final sample are summarized in Table 2. Respondents were fairly evenly represented by men and women and reflected a broad age range: between ages 50 and 93. Respondents were well educated and relatively well-off financially. In comparison with Current Population Survey (CPS) 2011 data for the population ages 60 and over, our specific sample is more highly educated and underrepresents racial and ethnic minority populations.¹⁹

Study participants reported considerable experience with chronic knee pain and knee osteoarthritis. More than 40% of respondents reported experiencing chronic knee pain, which is comparable to data from the 2011 Gallup-Healthways Well-Being Index.²⁰ Roughly half of these have been diagnosed with osteoarthritis in at least 1 knee, and 8% of our participants have had full knee replacement surgery. Two thirds of the sample reported having close relatives or friends diagnosed with knee osteoarthritis, and about 50% reported having close relatives or friends who have had full knee replacement surgery.

Data Analysis

The unit of analysis in this study is the single vignette observation. Because each respondent was

*The RAND American Life Panel draws its members from multiple survey programs, which makes the calculation of exact panel recruitment rate difficult. Under appropriate assumptions, approximate recruitment rates vary for the different sources between 18% and 29%.

Table 2 Descriptive Statistics for the Study Sample

Variables		Percentage
Gender	Male	43.4
Age (median = 59 years)	50–59	50.1
	60–69	33.7
	70 or older	16.3
	<\$25,000	23.4
Income	≥\$25,000 and <\$50,000	27.3
	≥\$50,000 and <\$75,000	16.1
	≥\$75,000	33.2
	Married or living with a partner	58.2
Living status	Retired	34.3
Employment status	Working	45.2
	Unemployed, disabled, and other	20.4
Education	High school or less	23.9
	At most bachelor's degree	59.2
	Postgraduate	16.9
Ethnicity	Non-Hispanic white	82.7
Respondents with chronic knee pain		42.1
Respondents with knee osteoarthritis		21.2
Respondents with full knee replacement surgery		8.2
Friends/relatives with knee osteoarthritis		65.8
Friends/relatives with full knee replacement surgery		52.5

presented with 3 vignette persons, we obtained up to 3 vignette observations from each of the 1616 respondents. With 3 instances of item nonresponse, our vignette sample consists of a total of 4845 vignette observations.

For each vignette observation, we identified whether the respondent recommended surgery for this specific vignette person. Our dependent variable “surgery recommendation” thus takes the value 1 if surgery was recommended and 0 if surgery was not recommended.

We describe the association between dimension levels and surgery recommendations as the proportion of vignettes presenting each level of a dimension in which respondents recommend surgery. We fitted logistic models for multivariate analysis of the contributions of several independent forms of information on surgery recommendations and present effect sizes as estimated odds ratios.

We conducted 2 independent analyses. The first analysis uses observations from the first and second vignettes to examine the effects of the basic dimensions, second opinions, and person-specific outcome forecasts. The second analysis uses observations obtained from the third vignette where patient testimonials could be available and person-specific outcome forecasts were optional. This analysis examines the relative impact of all 3 supplementary

information features and examines whether the addition of patient testimonials modifies the effects of the other dimensions. In both analyses, the estimated logit models control for random assignment to a text treatment (21% of the final sample), vignette order effects, and vignette-person-specific constants to account for the 6 different male and female vignette persons. We also control for the following respondent characteristics: gender, age, marital status, education, respondent's labor force status, ethnicity, and yearly household income. Additionally, we control for the following respondent health characteristics: whether the respondent has chronic knee pain, whether the respondent was diagnosed with knee osteoarthritis, whether the respondent has had full knee replacement surgery, whether the respondent has friends/relatives who have been diagnosed with knee osteoarthritis, and whether the respondent has friends/relatives who have had full knee replacement surgery. In the first analysis, standard errors are clustered at the individual level to account for the fact that each respondent was observed twice; in the second analysis, Huber-White robust standard errors are used.

The dimensions of our design are included in the regression models via effects coding. When viewing information is optional, we also include a dummy variable that captures when the dimension was offered but not chosen. For all 3 supplementary

Table 3 Surgery Recommendation Patterns

Sample	Percentage Recommending Surgery
Full sample	51.7
Panel A: Dimensions available at all vignettes ($N = 4845$). Sample split by:	
Pain level	
High pain	61.3
Low pain	41.5
Opportunity cost level	
High opportunity costs	47.0
Low opportunity costs	56.5
Employment status	
Retired	54.4
Employed	48.9
Surgeon's recommendation level	
Positive surgeon's recommendation	66.2
Neutral surgeon's recommendation	47.0
Second opinion level seen	
Strong recommendation	66.7
Substantial reservation	32.4
Second opinion not chosen	59.6
Second opinion not available	57.5
Person-specific outcome forecast level seen	
Forecast above average	64.9
Forecast below average	39.6
Numeric forecast format	54.1
Graphic forecast format	53.9
Verbal forecast format	48.2
Forecast not available	50.4
Forecast not chosen (third vignette)	53.3
Panel B: Patient testimonial dimension ($N = 1614$). Sample split by:	
Patient testimonial combination seen	
Full sample	54.5
1 positive testimonial	62.1
1 negative testimonial	47.5
2 positive testimonials	66.4
2 negative testimonials	41.4
Mixed message testimonials	55.2
Testimonials not available	53.7
Second testimonial not chosen	55.6

information dimensions, the reference category consists of the observations where the respective feature was not offered.

All P values in the multivariate analysis refer to tests for equality of the strength of the estimated influences using Wald tests. We performed analyses using Stata software (StataCorp, College Station, Texas).

FINDINGS

Distribution of Surgery Recommendations

For 52% of the vignette observations, participants recommended surgery for the vignette person. Table 3 shows the percentage of respondents who recommended surgery for the different dimension levels.

Panel A shows that the percentage of surgery recommendations for vignettes with high pain was almost 20 percentage points higher than for vignettes with low pain. There were similar differences in the percentages of surgery recommendations based on surgeons' recommendations. The differences based on opportunity costs and employment status were less pronounced but show the expected sign and are still significant if simple t tests for the difference of means are used.

Differences in the percentage of surgery recommendations for second opinions and the outcome forecasts are as strong as for the basic dimensions. For subsamples where a "strong recommendation" second opinion or an above-average outcome forecast was shown, respondents' recommendations for surgery are at least 25 percentage points higher than for subsamples with the "substantial reservations" second opinion or "below average likelihood of success" outcome forecast.

Panel B presents the distribution of surgery recommendations for only the third vignette observations, depending on which testimonial combination was seen and whether the testimonials were available. The subsamples where 1 or 2 positive testimonials were seen have higher percentages of surgery recommendations than the subsamples where 1 or 2 negative testimonials were seen. The subsample where mixed (positive and negative) testimonials were seen does not differ substantially from the subsample where no testimonial was offered.

Multivariate Analysis

Basic dimensions, second opinions and person-specific outcome forecasts. Model 1 in Table 4 uses 3231 observations from the first and second vignette observations. Due to the panel nature of the data, the data points in Model 1 in Table 4 are not independent. To address that concern, we refitted the model with random effects. The results were not substantively altered and are available from the authors upon request.

Model 1 shows that the effects of our basic dimensions are as expected. If a vignette person was

Table 4 Predictors of Respondent's Decisions Regarding Knee Replacement Surgery

	1. Model (n = 3231)			2. Model (n = 1614)		
	Odds Ratios	P	95% CI	Odds Ratios	P	95% CI
The influence of basic dimension levels						
Pain						
Low	Ref.			Ref.		
High	2.746	0.000	2.398–3.144	2.835	0.000	2.247–3.577
Opportunity costs						
Low	Ref.			Ref.		
High	0.585	0.000	0.513–0.667	0.606	0.000	0.481–0.763
Employment status						
Employed	Ref.			Ref.		
Retired	1.307	0.000	1.148–1.488	1.329	0.015	1.056–1.673
Surgeon's recommendation						
Neutral	Ref.			Ref.		
Positive	2.774	0.000	2.366–3.252	2.569	0.000	1.959–3.368
The influence of supplementary information levels						
Second opinion						
Not available	Ref.			Ref.		
Not chosen	1.272	0.089	0.964–1.679	1.371	0.173	0.871–2.158
Strong recommendation	1.835	0.000	1.547–2.175	1.885	0.000	1.380–2.577
Substantial reservation	0.321	0.000	0.269–0.382	0.346	0.000	0.257–0.467
Person-specific outcome forecast						
Not available	Ref.			Ref.		
Not chosen				0.885	0.632	0.537–1.458
Above average	1.831	0.000	1.498–2.238	1.707	0.004	1.183–2.463
Below average	0.515	0.000	0.420–0.630	0.413	0.000	0.288–0.592
Verbal format	Ref.			Ref.		
Numeric format	1.304	0.006	1.078–1.578	1.559	0.016	1.086–2.237
Graphic format	1.364	0.002	1.122–1.657	1.584	0.010	1.115–2.249
Patient testimonials						
Not available	Ref.			Ref.		
2 positive				2.280	0.001	1.417–3.670
2 negative				0.506	0.003	0.324–0.789
Mixed message				1.173	0.291	0.872–1.576
1 positive				1.410	0.092	0.946–2.100
1 negative				0.751	0.217	0.476–1.184
The influence of respondent's characteristics						
Age						
Gender	1.004	0.525	0.992–1.016	1.009	0.349	0.991–1.027
Female	Ref.			Ref.		
Male	1.060	0.465	0.906–1.240	1.241	0.086	0.970–1.588
Living status						

(continued)

Table 4 (continued)

	1. Model (n = 3231)			2. Model (n = 1614)		
	Odds Ratios	P	95% CI	Odds Ratios	P	95% CI
Living alone	Ref.			Ref.		
Married or living with a partner	0.991	0.921	0.832–1.181	1.079	0.575	0.827–1.407
Education						
High school or less	Ref.			Ref.		
Bachelor's degree or lower	0.854	0.116	0.702–1.040	0.810	0.156	0.606–1.084
Master's degree or more	0.716	0.015	0.547–0.938	0.649	0.032	0.437–0.963
Income						
<\$25,000	Ref.			Ref.		
\$25,000 to <\$50,000	1.035	0.768	0.822–1.305	1.126	0.509	0.792–1.600
\$50,000 to <\$75,000	0.963	0.781	0.736–1.259	1.181	0.424	0.785–1.776
≥\$75,000	1.091	0.501	0.847–1.405	1.223	0.311	0.829–1.804
Employment status						
Unemployed, disabled, and other	Ref.			Ref.		
Retired	0.914	0.473	0.713–1.170	0.853	0.433	0.574–1.269
Working	0.951	0.654	0.764–1.184	1.101	0.581	0.782–1.550
Ethnicity						
Other	Ref.			Ref.		
Non-Hispanic white	0.825	0.063	0.674–1.011	0.685	0.021	0.496–0.945
Health characteristics						
Respondent has chronic knee pain	0.773	0.007	0.640–0.932	0.846	0.269	0.629–1.138
Respondent has knee osteoarthritis	0.922	0.506	0.725–1.172	0.851	0.387	0.590–1.227
Respondent has had full knee replacement surgery	2.115	0.000	1.534–2.917	2.382	0.000	1.498–3.786
Respondent's close friends/relatives have knee osteoarthritis	0.840	0.137	0.668–1.057	0.854	0.402	0.590–1.236
Respondent's close friends/relatives have had full knee replacement surgery	1.513	0.000	1.216–1.883	1.399	0.063	0.983–1.993
Text control						
Video mode	Ref.			Ref.		
Text mode	0.934	0.467	0.778–1.122	1.009	0.951	0.757–1.345
Vignette fixed effects						
Joe	0.683	0.350	0.307–1.520	0.454	0.202	0.135–1.528
Bill	0.488	0.081	0.218–1.092	0.291	0.045	0.087–0.970
Susan	0.752	0.488	0.335–1.685	0.557	0.344	0.165–1.875
Ralph	1.010	0.982	0.450–2.265	0.750	0.648	0.218–2.577
Rosa	0.593	0.204	0.265–1.328	0.360	0.101	0.106–1.222
Betty	0.647	0.291	0.288–1.451	0.479	0.236	0.142–1.618
Log pseudo-likelihood		-2717.3277				-886.999

Note: CI = confidence interval; Ref. = reference category.

characterized by a high level of pain, the odds of a surgery recommendation were 2.75 times higher than for a vignette person with low pain. Similarly, a positive surgeon's recommendation made the odds of a surgery recommendation 2.77 times higher than a neutral surgeon's recommendation. Vignette persons with high opportunity costs had lower odds of receiving a surgery recommendation than vignette persons with low opportunity costs, and retired vignette persons have somewhat higher odds to receive a surgery recommendation than employed vignette persons.

Second opinions and outcome forecasts affected choices in an equally strong way. If a vignette person was presented with a second opinion that strongly recommended surgery, this vignette person's odds of getting a surgery recommendation were 1.84 times higher than for vignette persons where no second opinion was offered. When a vignette person was presented with a second opinion that raised substantial reservations, the odds of receiving a surgery recommendation were lower than for vignette observations where no second opinion was offered.

Vignettes with above-average outcome forecasts had 1.83 times higher odds of receiving a surgery recommendation, and vignettes presented with below-average outcome forecasts had lower odds of receiving surgery recommendations, than vignette persons where no outcome forecasts were provided.

The effects of patient testimonials. Model 2 in Table 4 only uses the 1614 observations obtained from the third vignette. It extends Model 1 by adding the effects of viewing a specific testimonial version and by controlling for the fact that reviewing outcome forecasts was optional. In this model, we obtained odds ratios on the basic dimensions, second opinions, and outcome forecasts that are comparable and statistically undistinguishable to the odds ratios of Model 1. However, in Model 2, we additionally found strong independent impacts of 2 testimonial versions on recommendations. Viewing 2 positive testimonials significantly increases the odds of a surgery recommendation by 2.28; viewing 2 negative testimonials significantly decreases the odds. Viewing only 1 testimonial or a testimonial combination with a mixed message did not significantly change the odds of receiving a surgery recommendation. Controlling for previous choices did not alter these results.

Based on this model, the effect sizes of a strong recommendation second opinion, an above-average outcome forecast, and 2 positive patient testimonials are not significantly different from each other (P values

range between 0.3324 and 0.6920). The effectiveness of a substantial reservations second opinion, a below-average outcome forecast, and 2 negative patient testimonials are also not significantly different from each other (P values range between 0.1611 and 0.4693). The effect of strongly recommending second opinions cannot be distinguished from the effect of a positive surgeon's recommendation. Pain evolved as a stronger driver of choice, its effect being significantly larger than the effect of a recommending second opinion ($P = 0.0355$), the effect of an above average outcome forecast ($P = 0.0182$), and the effect of being retired ($P < 0.0001$).

DISCUSSION

In this stated choice experiment, both outcome forecasts and testimonials influenced patient decisions regarding full knee replacement surgery. The effects of outcome forecasts and testimonials were similar to the effects of opportunity costs and second opinions but not as strong as the effects of pain. Testimonials influenced decisions only if the 2 messages about patient experiences with surgery were consistently positive or negative.

Our results regarding the effects of patient testimonials add to the growing body of literature that cautions about their use in decision support strategies because of their potential to bias decisions,³ especially if they are not representative.⁴ In contrast to some of the studies reviewed by Winterbottom and others,³ the introduction of testimonials to the decision-making process did not crowd out the effects of the statistical information contained in the person-specific outcome forecasts in our study. This may be due to the fact that testimonials were not clearly given a predominant role in the choice task in our study. In our research protocol, respondents were carefully briefed through a physician's lecture on knee osteoarthritis, treatment options, full knee replacement surgery, and the risks and side effects associated with this surgery and rehabilitation. They were presented with a good deal of information that described each vignette person's personal situation including the symptoms, treatment history, pain level, employment status, and opportunity costs that would be encountered in the case of surgery. Furthermore, they were presented with the orthopedic surgeon's recommendation as well as the recommendation by a second physician, described as a specialist in treating knee osteoarthritis. They reviewed person-specific outcome forecasts that

were presented as based on a large study of full knee replacement surgery patients funded by the US Department of Health and Human Services. However, it is worrying that despite this person-specific, professional, and objective information, the subjective testimonial of 2 lay people's negative or positive experiences with full knee replacement surgery still had an effect that was similar in magnitude to the effects of the most influential dimensions. Thus, our findings reinforce the importance of the concern raised in other studies about the potential for testimonials to have adverse effects on the quality of patient decision making. Additionally, the fact that the vast majority of respondents chose to view testimonials when offered (74%, see Appendix B online) is evidence that the public finds testimonials attractive.

With respect to the impact of statistical outcome data, our findings pertaining to the influence of the person-specific outcome forecasts show greater potential than previous studies for statistical data to stand up to testimonial messages. In previous studies reported by Greene and others²¹ and Betsch and others,⁴ in which the relative strengths of statistical information and testimonials were compared, testimonials were found to be more powerful. In our study, statistical outcome data and testimonials had comparable effects. It is possible that the relatively strong impact of statistical outcome forecasts in our study can be attributed to the fact that they were patient-specific.

The strong impact of pain on decision making is in line with other literature that has examined the influence of pain on the decision to undergo hip or knee replacement surgery or on the willingness to consider surgery²²⁻²⁴ as well as its relevance in the design of surgery waiting lists.^{25,26}

One limitation that may affect the generalizability of this study is the use of a discrete choice experiment, a stated-preference approach with a general population sample. We used hypothetical scenarios to elicit choices. Our premise is that these choices approximate the behavior of patients in real situations. There is longstanding debate about the degree to which results obtained from stated-preference approaches can be generalized to real-world decisions.²⁷⁻³¹ Despite the uncertainties about the potential for generalizing from stated-preference studies, these designs make it possible to examine the effects of options that could not be studied in research on actual patients because of ethical concerns.

Another limitation concerns the findings regarding the testimonial effects. In anticipation of the sequence with which patients might obtain information, the testimonials were the last piece of

information provided to respondents; thus, there is the possibility of recency effects.³² However, we offered respondents summaries of the other information in vignettes after they viewed the testimonials but before they made their recommendation. We anticipated that the summaries would remind respondents of vignette details and thus be as available as the testimonials at the moment of decision making. The question whether testimonial effects were possibly augmented by recency effects could be addressed through further experimentation by introducing testimonials earlier.

Our research did not address questions about the effects of general outcome forecasts that are now widely used in decision guides. Our findings suggest that patients would be influenced by forecasts specific to their circumstances. With our study, we can only hypothesize that these specific forecasts might be more influential. Future experimental research might systematically compare the effects of patient-specific and general-outcome forecasts.

We limited respondents to 2 testimonials. The fact that we found that the number of testimonials made a difference (when they were in the same direction) invites questions about the effects of viewing a larger number of testimonials. However, presenting more hypothetical choice tasks could also induce fatigue, and there arises an interesting question about the optimal number of such tasks, given the tradeoff between obtaining more information and that information potentially becoming less reliable.

Our findings have implications for the manner in which testimonials are used in decision guides for patients. Because testimonials can have a big impact on patients and can serve multiple purposes, such as to inform and comfort patients,³³ the inclusion of testimonials helps make guides attractive to patients. However, because patients can be unduly influenced by outcomes in a small unrepresentative sample of testimonials, patient guides should include testimonials with a balance of more favorable and less favorable outcomes in an effort to minimize the effect on decision making. Developers of decision guides may also use testimonials primarily to call attention to useful patient strategies in considering treatment alternatives. In doing so, they can minimize the attention given to outcomes described in the testimonials. In this way, testimonials can be used to engage patients fully in the decision process with a reduced risk of excessively influencing the choices that patients make.

Furthermore, our findings are promising with respect to the use of person-specific outcome forecasts in decision support strategies. Not only did

they have a strong effect on decision making, but they were also highly demanded by respondents when offered (90%, see Appendix B online). The fact that in our study, these forecasts were patient specific and presented as coming from a credible source has most likely contributed to this high demand and the strong effects we observed. While in many patient consultations subjective predictions are presented to the patient by the doctor, we hypothesize that patients are likely to give more weight to predictions that are based on outcome research and not simply on the subjective probability of the treating physician. Additionally, if outcome forecasts are personalized (i.e., produced using personal and health-related information that is specific to the respective patient), they may contain much more relevant information to the patient and may thus be given more weight in the decision-making process than generalized forecasts. Another side effect may be that the understanding of what these statistics mean for the specific patient may be easier since some of his or her important personal characteristics are already taken into account and do not have to be used to interpret a general outcome forecast. However, since communicating statistical information effectively is difficult, we believe that the recent research calling for better training of physicians in this area³⁴ should be taken seriously.

In conclusion, our results call for a cautious use of patient testimonials in decision support strategies, perhaps with a modified emphasis of engaging patient interest and exemplifying possible decision-making strategies. Furthermore, we demonstrated that person-specific outcome forecasts strongly influence decisions and could thus be a useful addition to decision aids as well as medical consultations in general.

REFERENCES

1. International Patient Decision Aids Standards Collaboration (IPDAS). Available from: URL: <http://ipdas.ohri.ca/index.html>
2. Bekker H, Winterbottom A, Fowler J, et al. 2012 Updated chapter E: using personal stories. In: Volk R, Llewellyn-Thomas H, eds. Update of the International Patient Decision Aids Standards (IPDAS) Collaboration Background Document. Available from: URL: <http://ipdas.ohri.ca/IPDAS-Chapter-E.pdf>
3. Winterbottom A, Bekker HL, Conner M, Mooney A. Does narrative information bias individual's decision making? A systematic review. *Soc Sci Med*. 2008;67:2079–88.
4. Betsch C, Ulshofer C, Renkewitz F, Betsch T. The influence of narrative v. statistical information on perceiving vaccination risks. *Med Decis Making*. 2011;31(5):742–53.
5. Chaiken S, Wood W, Eagly AH. Principles of persuasion. In: Higgins ET, Kruglanski AW, eds. *Social Psychology: Handbook of Basic Principles*. New York: Guilford Press; 1996:702–42.
6. March JG. *A Primer on Decision Making: How Decisions Happen*. New York: The Free Press; 1994.
7. Rossi P, Anderson A. The factorial survey approach: an introduction. In Rossi P, Nock S, eds. *Measuring Social Judgments: The Factorial Survey Approach*. Beverly Hills (CA): Sage; 1982.
8. Louviere J, Hensher D, Swait J. *Stated Choice Methods: Analysis and Application*. New York: Cambridge University Press; 2000.
9. Johnson FR, Kanninen B, Bingham M, Ozdemir S. Experimental designs for stated choice studies. In: Kanninen BJ, ed. *Valuing Environmental Amenities Using Stated Choice Studies*. Dordrecht (Netherlands): Springer; 2007:159–92.
10. Caro F, Ho T, McFadden D, Gottlieb A, Yee C, Winter J. Using the internet to administer more realistic vignette experiments. *Soc Sci Comput Rev*. 2012;30(2):184–201.
11. Health Dialog. *Treatment Choices for Knee Osteoarthritis*. Boston (MA): Foundation for Informed Medical Decision Making; 2007.
12. Gurmankin A, Baron J, Hershey J, Ubel P. The role of physicians recommendations in medical treatment decisions. *Med Decis Making*. 2002;22(3):262–71.
13. Moumjid N, Gafni A, Bremond A, Carrere M. Seeking a second opinion: do patients need a second opinion when practice guidelines exist? *Health Policy*. 2007;80(1):43A–50.
14. Van Dalen I, Groothoff J, Stewart R, Spreeuwenberg P, Groenewegen P, van Horn J. Motives for seeking a second opinion in orthopedic surgery. *J Health Serv Res Policy*. 2001;6(4):195–201.
15. RAND American Life Panel. Available from: URL: <https://mmicdata.rand.org/alp/index.php?page=panel>
16. Chang L, Krosnick JA. National surveys via RDD telephone interviewing versus the Internet: comparing sample representativeness and response quality. *Public Opin Q*. 2009;73:641–78.
17. Yeager DS, Krosnick JA, Chang L, et al. Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opin Q*. 2011;75(4):709–47.
18. The American Association for Public Opinion Research. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 7th edition. AAPOR, Deerfield (IL): AAPOR; 2011.
19. US Census Bureau American FactFinder. Population 60 years and over in the United States. 2011 American Community Survey 1-year estimates. Available from: URL: <http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml>
20. Gallup-Healthways Well-Being Index. Available from: URL: <http://www.gallup.com/poll/154169/chronic-pain-rates-shoot-until-americans-reach-late-50s.aspx>
21. Greene K, Campo S, Banerjee SC. Comparing normative, anecdotal, and statistical risk evidence to discourage tanning bed use. *Commun Q*. 2010;58(2):111–32.
22. McHugh GA, Luker KA. Influences on individuals with osteoarthritis in deciding to undergo a hip or knee joint replacement: a qualitative study. *Disabil Rehabil*. 2009;31(15):1257–66.
23. Toye FM, Barlow J, Wright C, Lamb SE. Personal meanings in the construction of need for total knee replacement surgery. *Soc Sci Med*. 2006;63(1):43–53.

24. Hawker GA, Wright JG, Badley EM, Coyte PC. Perceptions of, and willingness to consider, total joint arthroplasty in a population-based cohort of individuals with disabling hip and knee arthritis. *Arthritis Care Res.* 2004;51(4):635–41.
25. Sampietro-Colom L, Espallargues M, Rodriguez E, et al. Wide social participation in prioritizing patients on waiting lists for joint replacement: a conjoint analysis. *Med Decis Making.* 2008;28(4):554–66.
26. Allepuz A, Espallargues M, Moharra M, Comas M, Pons J. Prioritisation of patients on waiting lists for hip and knee arthroplasties and cataract surgery: instruments validation. *BMC Health Serv Res.* 2008;8:76.
27. Diamond P, Hausman J. Contingent valuation: is some number better than no number? *J Econ Perspect.* 1994;8(4):45–64.
28. Hensher D. Hypothetical bias, choice experiments and willingness to pay. Working Paper ITLS-WP-09-01. 2009. p 1–29. Available from: URL: http://sydney.edu.au/business/__data/assets/pdf_file/0009/25479/itls_wp_09_01.pdf.
29. Mark T, Swait J. Using stated preference and revealed preference modeling to evaluate prescribing decisions. *Health Econ.* 2004;13:563–73.
30. Harrison G. Hypothetical bias over uncertain outcomes. In: List JA, ed. *Using Experimental Methods in Environmental and Resource Economics.* Northampton (MA): Edward Elgar; 2006. p 41–69.
31. Kesternich I, Heiss F, McFadden D, Winter J. Suit the action to the word, the word to the action: hypothetical choices and real decisions in Medicare Part D. *J Health Econ.* 2013;32(6):1313–24.
32. Baddeley AD. *Human Memory: Theory and Practice.* Needham Heights (MA): Allyn & Bacon; 1990.
33. Shaffer V, Zikmund-Fisher BJ. All stories are not alike: a purpose-, content-, and valence-based taxonomy of patient narratives in decision aids. *Med Decis Making.* 2013;33(1):4–13.
34. Gaissmaier W, Anderson BL, Schulkin J. How do physicians provide statistical information about antidepressants to hypothetical patients? *Med Decis Making.* 2014;34(2):206–15.