ORIGINAL PAPER

The elicitation of patient and physician preferences for calculating consumer-based composite measures on hospital report cards: results of two discrete choice experiments

Martin Emmert[1] · Stefan Rohrbacher[1] · Florian Meier[2] · Laura Heppe[3] · Cordula Drach[3] · Anja Schindler[4] ·
Uwe Sander[4] · Christiane Patzelt[4] · Cornelia Frömke[4] · Oliver Schöffski[3] · Michael Lauerer[1]

Abstract

Purpose The calculation of aggregated composite measures is a widely used strategy to reduce the amount of data on hospital report cards. Therefore, this study aims to elicit and compare preferences of both patients as well as referring physicians regarding publicly available hospital quality information

Methods Based on systematic literature reviews as well as qualitative analysis, two discrete choice experiments (DCEs) were applied to elicit patients' and referring physicians' preferences. The DCEs were conducted using a fractional factorial design. Statistical data analysis was performed using multinomial logit models

Results Apart from five identical attributes, one specific attribute was identified for each study group, respectively. Overall, 322 patients (mean age 68.99) and 187 referring physicians (mean age 53.60) were included. Our models displayed significant coefficients for all attributes ($p < 0.001$ each). Among patients, "Postoperative complication rate" (20.6%; level range of 1.164) was rated highest, followed by "Mobility at hospital discharge" (19.9%; level range of 1.127), and "The number of cases treated" (18.5%; level range of 1.045). In contrast, referring physicians valued most the "One-year revision surgery rate" (30.4%; level range of 1.989), followed by "The number of cases treated" (21.0%; level range of 1.372), and "Postoperative complication rate" (17.2%; level range of 1.123)

Conclusion We determined considerable differences between both study groups when calculating the relative value of publicly available hospital quality information. This may have an impact when calculating aggregated composite measures based on consumer-based weighting.

Keywords Public reporting · Hospital choice · Discrete choice experiment · Composite measures · Hospital report cards

JEL Classification I12

✉ Martin Emmert
Martin.Emmert@uni-bayreuth.de

1 Faculty of Law, Business and Economics, Institute for Healthcare Management and Health Sciences, University of Bayreuth, Prieserstraße 2, 95444 Bayreuth, Germany

2 Department of Management and Economics, SRH Wilhelm Löhe University of Applied Sciences, 90763 Fürth, Germany

3 School of Business and Economics, Chair of Health Care Management, Friedrich-Alexander-University of Erlangen-Nuremberg, Lange Gasse 20, 90403 Nuremberg, Germany

4 Department of Information and Communication, Faculty for Media, Information and Design, University of Applied Sciences and Arts, Hannover, Germany

Introduction

Public reporting of performance information encompasses data, publicly available or available to a broad audience free of charge or at a nominal cost, about a health-care structure, process, or outcome at any provider level (e.g., hospitals) [1]. The overall aim of public reporting is to improve health-care quality by both stimulating quality improvement on the provider level ("Improvement Through Changes in Care") and also by helping patients, referring physicians, and other consumers select the "right" provider ("Improvement Through Selection") [2, 3]. Therefore, hospital report cards (HRCs) publicly display quality-related information about hospitals and enable hospital comparisons [4].

However, studies have shown heterogeneous results regarding the impact of HRCs on the choice behavior of consumers [5–10]. To increase the uptake of HRCs of consumers, HRCs should present the information that consumers value most when making hospital choices [11]. Besides, it should be attempted to limit the richness of consumers' choice sets, i.e., the number of hospitals as well as the variety of quality measures [12].

In this context, one promising strategy is to provide customized information that both reflects the preferences of consumers and reduces the amount of data by aggregating individual performance measures into summary scores (i.e., composite measures) [11, 13–17]. For example, a recent study has shown that the introduction of summary measures on Nursing Home Compare—a web-based guide detailing quality of care at over 17,000 Medicare- or Medicaid-certified nursing homes in the United States—was associated with a significant change in consumer demand for low- and high-scoring facilities [18]. So far, composite measures have been published in several countries such as the United States, the UK, and Germany (e.g., Hospital Compare, Nursing Home Compare, NHS Choices, US News Best Hospital, AOK Hospital Navigator) [17–19].

There are different approaches to calculate aggregated composite measures such as opportunity-based weights, numerator-based weights, all-or-none measures, expert panels, and consumer-based weighting [16]. As shown, consumer-based weighting aligns mostly closely with patient-centered care and should, therefore, be prioritized [16]; for this purpose, it is crucial to learn more about consumers' preferences when choosing a hospital. So far, current composite measure approaches do not consider consumer preferences in more detail when calculating summary scores. For example, the overall star rating system for hospitals for the Centers for Medicare and Medicaid Services (CMS) in the United States—which is publicly displayed on Hospital Compare—is calculated by taking the weighted average of the hospital's scores across five areas of quality (i.e., mortality, safety of care, readmission, patient experience, and timely and effective care). Therefore, the relative weight for mortality, safety, readmission, and patient experience is worth 22% each, and timely and effective care is worth 12% of the overall score [20].

In addition, most research has been conducted to investigate whether patients—as the main target group of HRCs—use publicly reported quality information to search for and select hospitals [21–23]. However, less information is available regarding whether publicly available quality information plays a role from the referring physicians' perspective [24, 25]. This seems to be surprising since most patients trust their referring physicians' recommendation regarding what hospital to choose [25, 26]. Therefore, referring physicians should be regarded as another major target group of HRCs so as to direct patients to well-performing hospitals and to increase the impact of public reporting [24, 27].

In this context, the present study aims to identify the most important publicly reported hospital quality information for hospital choice for elective hip replacement surgery and to determine their relative importance among both patients as well as referring physicians. We focused on elective hip replacement surgery for the following reasons: first, we aimed to address one procedure that is both standard for public reporting initiatives as well as included in German hospital quality assurance initiatives [28, 29]. Second, non-acute procedures should be prioritized since public reporting aims to support patients, referring physicians, and other consumers to select the "right" provider. [2, 3] Third, elective hip replacement surgery is one of the most frequently conducted procedures in Germany; for example, 160,910 surgeries were performed in 1,254 hospitals in 2020 [30]. The results might help us learn how to weight different quality measures when calculating composite measures. In more detail, we sought to address the following three questions: (1) what is the essential publicly available hospital quality information for choosing a hospital for elective hip replacement surgery from the perspective of patients and referring physicians? (2) How do patients and referring physicians' rate different publicly available hospital quality information? (3) What relative importance do both patients and referring physicians assign to different quality information?

Materials and methods

This study used a mixed methods approach. After performing a systematic literature review, we conducted qualitative research methods (i.e., semi-structured interviews) to identify and select the most important quality measures for choosing a hospital for hip replacement surgery from the perspective of both patients and referring physicians. Based on this, two separate discrete choice experiments (DCEs) were developed and performed to elicit and compare patients' and referring physicians' preferences for all relevant hospital choice characteristics and to determine the relative value of each quality measure. DCEs are increasingly used in the health-care context to inform on consumer preferences for health-care services such as hospital choice [31]. DCEs are a stated preference method that use (survey) data to systematically elicit individuals' preferences based on a series of hypothetical choice scenarios (termed choice sets) [32]. Two key economic theories (i.e., Random Utility Theory and Lancaster's Theory [33, 34]) suggest that respondents choose the option from each choice set which provides them with the most satisfaction or ''utility'' [32]. Based on the decisions, DCEs can help understand which characteristics (termed attributes) are preferred by

consumers and determine the relative value of each attribute. [32, 35]

Study design

The design and analysis of both DCEs were based on standardized research practices for undertaking conjoint analysis of the ISPOR Conjoint Analysis Good Research Practices Task Force [36–38]. Ethics approval for both studies was obtained from the Friedrich-Alexander-University Erlangen-Nuremberg Ethics Board (196_19 B). Informed consent was obtained from each study participant.

Systematic search procedure

First, we conducted two systematic search procedures on Medline (via PubMed) and the Cochrane Library to identify studies which aimed to identify relevant criteria for hospital choice for hip replacement surgery from the perspective of both patients and referring physicians. The searches were carried out in September 2020 and aimed to identify English and German language literature published since 2010. In addition, reference lists of identified research articles were screened for further articles. The reviews were complied with the Guideline from the Cochrane Collaboration. [39]

The search strategy addressing the *perspective of patients* was segmented into three components. The first component referred to hospitals (e.g., hospital, clinic), the second component referred to choice (e.g., choice, selection), and the third component addressed hip replacement surgery (e.g., hip, coxarthrosis). As a result, 5,510 potentially relevant papers were identified. After eliminating duplicates and judging titles and abstracts in a first step as well as full papers in a second step, ten studies were considered relevant [40–49]. In sum, 73 individual criteria were derived from the 10 studies. We added four further quality measures from the German hospital quality report administered by the Institute for Quality Assurance and Transparency in Healthcare (IQTIG). This led to a final sample of 77 quality measures which were then merged into 23 more general measures (e.g., the following three individual criteria "28-day mortality rate (%)", "28-day mortality rate following discharge", and "Postoperative mortality" were merged into the category "Mortality rate"). Afterward, we excluded ten criteria which are not available in the German hospital sector for public reporting purposes (e.g., waiting time, 28-day emergency readmission rate), so that 13 criteria remained for the next step (See Supplemental Material 1).

The search strategy addressing the *perspective of referring physicians* was based on previously published literature and slightly modified [50]. As a result, 2,246 potentially relevant papers were identified. After eliminating duplicates and judging titles and abstracts in a first step as well as full papers in a second step, 16 studies were considered relevant [25, 51–65]. In sum, 39 criteria were derived from those studies. Then we excluded 20 criteria that are not publicly available (see above) (e.g., waiting times, MRSA events), so that 19 criteria remained.

Qualitative steps (semi-structured interviews)

In total, we qualitatively surveyed (January 2021 to February 2021) 20 randomly selected hip replacement surgery patients (mean age 63.9 years, 55% female) from a German statutory health insurance who had undergone elective hip arthroplasty surgery within 3 years prior to answering the survey as well as 15 referring physicians (mean age 56.3 years, 20% female). Written informed consent was obtained from all participants. Respondents were sent a short survey through postal mail before conducting the semi-structured interviews to learn more about (the) past hospital choice(s) and the importance of the 13 resp. 19 criteria for choosing a hospital. After receiving the short survey, semi-structured interviews were conducted to explore the stated preferences, to determine the most relevant criteria as well as corresponding levels, to clarify the wording, and to evaluate the comprehensibility of hypothetical choice tasks for the DCE (see below). In particular, we evaluated the comprehensibility of the labels used for the attributes and their corresponding descriptions to assure the intended interpretation of the explanations. Based on all steps, five attributes were derived which were of major importance for choosing a hospital from the perspective of both patients and referring physicians which were as follows: the rate of confirmed diagnosis prior the surgery, the certification as an Endoprosthetics Center, the number of cases treated, the rate of postoperative complications, as well as the rate of mobility at hospital discharge. This was supplemented with the assessment of actions to prevent falls of patients for patients and the 1-year revision surgery rate for referring physicians, respectively (Table 1). Individuals who completed the qualitative study received 50 Euro.

Quantitative steps

The survey instrument

Both survey instruments were similar and consisted of four parts. First, we asked for general sociodemographic information (e.g., age, gender) before collecting information about the experience with HRCs when selecting hospitals in the second part. Second, respondents were presented with publicly available hospital quality information items (see above) as well as short descriptions and were asked to rate each item on a 1–5 scale (1 = not all important; 5 = very important). Furthermore, we asked the respondents to select the single

Table 1 Summary of attributes and levels used in the discrete choice experiments (DCE)

Attribute	*Description* and levels given in DCE questionnaire
Confirmed diagnosis rate[$,#]	*Hip replacement primary implantation with fulfilled indication criteria*
	Quality targets reached
	Quality targets not reached
	No assessment intended resp. results not (yet) available
EndoCert certificate[$,#]	*Hospitals which meet requirements of the EndoCert certification system*
	No certificate
	Certified EndoProstheticsCenter (EPZ)
	Certified EndoProstheticsCenter of Maximum Care (EPZmax)
The number of cases treated[$,#]	*The number of total elective hip replacement surgeries in a hospital*
	84 patients (below average)
	148 patients (average)
	230 patients (above average)
Postoperative complication rate[$,#]	*Postoperative specific complications after hip replacement surgery (e.g., infections, bleeding, thromboembolism)*
	Quality targets reached
	Quality targets not reached
	No assessment intended resp. results not (yet) available
Mobility at hospital discharge[$,#]	*Functional mobility status on discharge after hip replacement surgery*
	Quality targets reached
	Quality targets not reached
	No assessment intended resp. results not (yet) available
Prevention of falls measures[$]	*Hospitals should be active in minimizing fall risk among patients who have undergone hip replacement surgeries*
	Quality targets reached
	Quality targets not reached
	No assessment intended resp. results not (yet) available
1-year revision surgery rate[#]	*Percentage of patients who underwent revision surgery within one year after initial surgery*
	1.41% (better than average)
	2.48% (average)
	3.89% (lower than average)

[$]This attribute applies for patients

[#]This attribute applies for referring physicians

most important information item for the hospital choice. Afterward, all respondents were asked to respond to the DCE survey block. Both DCE experiments were designed using Sawtooth Software Lighthouse Studio 9.11.0.

Instead of ranking or rating different quality measures, as is done in traditional importance elicitation formats, DCEs perform a pairwise comparison of hypothetical alternatives (i.e., differently configured hospitals) and ask the participants to choose between them [66]. Both experiments were designed as a full profile design (i.e., each choice set included all six attributes) and were generated using the balanced overlap method so as to achieve standard errors below 0.05 for main effect utilities and 0.10 or smaller for interaction effects and the highest D-efficiency score [67, 68]. We created 10 versions (termed blocks) of the DCE, each containing 10 choice tasks, thus generating 100 unique choice tasks (fractional factorial design). Choice tasks were generated that maximized both balance (i.e., meaning each level appears with the same frequency) and orthogonality (i.e., meaning each pair of levels appears with the same frequency across each pair of attributes) [69]. The choice sets in both surveys were designed as forced-choice tasks; this means that respondents had to choose one of two hypothetical hospitals by making trade-offs between attributes and their levels. This method offers practical advantages such as closeness to reality as trade-off decisions are part of everyday life [66]. As stated above, hospitals differed in six attributes with three levels each. In DCE research, four to eight attributes per choice set are seen as appropriate [70, 71]. Figure 1 provides one hypothetical choice task as an example. The questionnaire was pilot tested for clarity and understanding
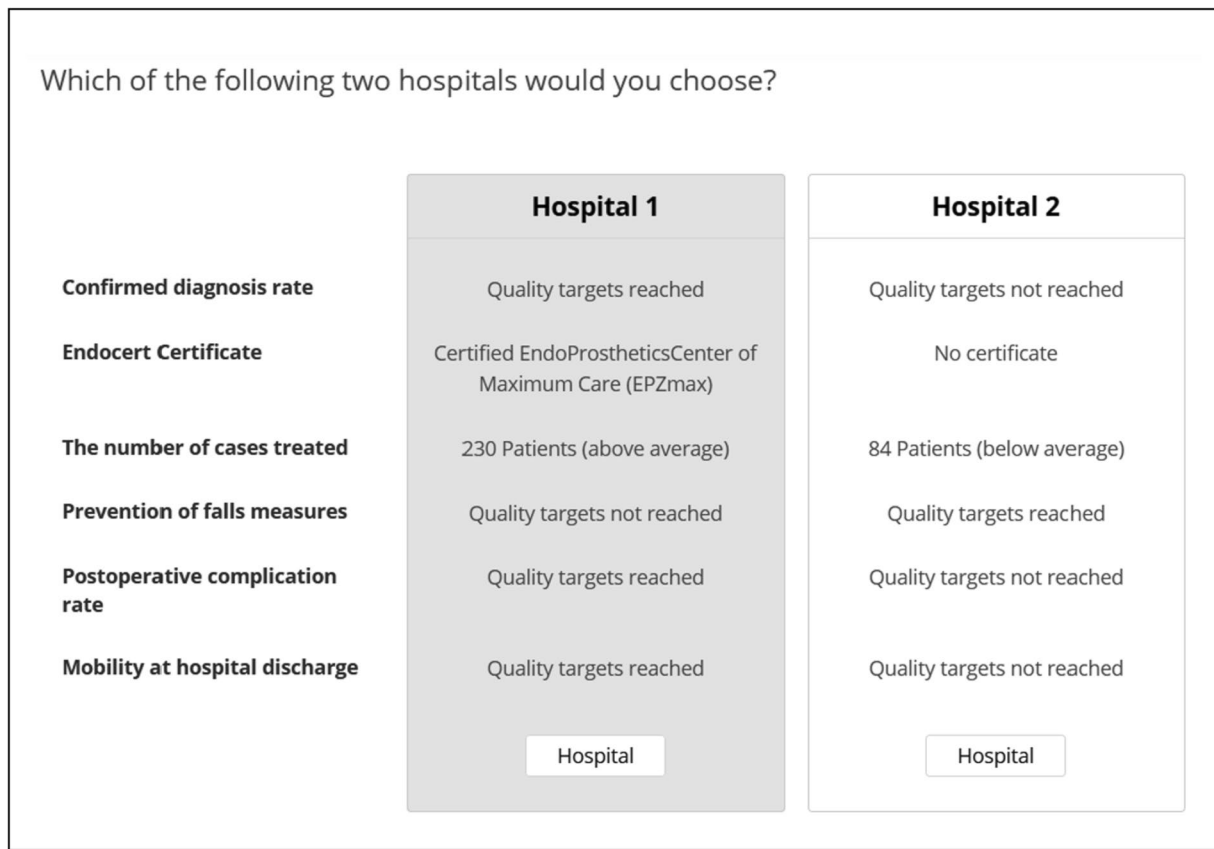
Fig. 1 Example choice task for patients

with 15 hip replacement patients as well as 15 referring physicians and slightly modified accordingly.

The survey sample

Final questionnaires were sent through postal mail to 1000 randomly selected people from a German statutory health insurance (AOK Lower Saxony) who had undergone elective hip arthroplasty surgery within 3 years prior to answering the survey. One reminder was sent out 2 weeks after the initial invitation. Furthermore, we purchased a database containing contact information (e.g., postal address, email) for orthopedists in the German outpatient sector from a commercial provider (ArztData AG). This database covers about 88% of all orthopedists in the German outpatient sector [72]. We used a sequential mixed-mode strategy to achieve high response rates. In a first step, 1,650 referring orthopedists were contacted via email, which contained a link to participate online (web-based survey). After 1 week, a first reminder was sent out. In a second step, the remaining orthopedists were contacted via postal mail and received a printed version of the survey; after 2 weeks, we sent out a second reminder. The surveys were conducted between March and June 2021 in the German language. As an incentive, respondents received a payment of approximately €50.

Sample size

Johnson's often-used rule-of-thumb calculates a sample of 75 participants for a DCE having our design specifications (i.e., ten choice tasks per respondent, two alternatives, three levels per attribute as maximum) [73]. This sample size was discussed as being the lowest limit for main effects estimation. However, we aimed at doubling this number (i.e., to include at least 150 participants) following more advanced recommendations for statistical robustness [74, 75].

Statistical methods

Analyses of general survey questions were performed using SPSS (IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp). Descriptive statistics (i.e., means for continuous variables, percentages for categorical variables) were used to examine the demographic and experience variables.

For analyzing the DCEs, we performed multinomial logit models using R (Version 4.2.2) and the corresponding

mlogit-package by Croissant (2020) [76]. We constructed the data sets in long format. Thus, the data set about patients included 6,440 observations (322 patients×10 choice tasks per subject×2 alternatives per choice set). Similarly, the data set for physicians comprised 3,740 observations (187 physicians×10 choice tasks per subject×2 alternatives per choice set). As for data encoding, we started from assigning each variable equidistant numerical values of scale {0; 1; 2} as described by Mühlbacher et al. [77]. For example, regarding the attribute "Confirmed diagnosis rate", we designated the value {0} to "Quality targets not reached", value {1} to "No assessment intended resp. results not (yet) available" (please note: due to the implicit uncertainty of this alternative, we assumed that the absence of results lay between the certain answers "targets reached" and "targets not reached"), and value {2} to "Quality targets reached". Afterward, we repeated the estimations treating the attributes as categorical variables. Both approaches provided qualitatively very similar results, whereby the estimates from categorical attributes could be shown to be more precise.

We employed effect coding by following standard guidelines [38, 78]. Therefore, we expressed every attribute as *n*-1 variables where n is the number of levels. Each variable corresponds to one level, while the omitted level serves as the reference. If the variable matches the level presented in the corresponding profile, we assigned value {1}. If the level in the corresponding profile equals another non-omitted level, we assigned value {0}. If the level in the corresponding profile is the omitted level, we assign value {− 1}. For example, the attribute "Confirmed diagnosis rate" was operationalized by three possible levels: "Quality targets reached", "Quality targets not reached", "No assessment intended resp. results not (yet) available". As reference level, we chose "No assessment intended resp. results not (yet) available". Hence, we constructed two variables; the first covered the level "Quality targets reached", the other one the level "Quality targets not reached". If the given occurrence equals the reference, we filled both variables with {-1}, otherwise we assigned values {1} or {0}. As reference categories, we assigned "Certified Endoprosthetics Center of Maximum Care (EPZmax)" for the attribute "Endocert Certificate", middle positions for the attributes "Number of cases treated" and "One-year revision surgery rate", and finally "No assessment intended resp. results not (yet) available" for all remaining attributes. We chose effect coding over dummy coding since with dummy coding, the parameter estimate for the (omitted) baseline category is equal to zero and cannot be recovered. Thus, the estimates of the other levels are only relative to the benchmark level. Instead, with effect coding, we can extract the parameter of the reference category from the negative sum of the included categories and the standard error from the covariances of the included categories. This enables us to compare all levels of the attributes against the corresponding mean value with the sign of the coefficients indicating a positive or negative impact compared to the mean of the attribute and the magnitude specifying the size of the effect [38, 78, 79]. In all approaches, we applied Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and log-likelihood value (LL) to identify the best fitting model. The relative importance of each attribute was determined based on each attribute's coefficient range. This means that we calculated the difference between the coefficients of the highest and lowest level of each attribute and expressed this as a share of the total range across all attributes.

Results

Sample characteristics

Overall, 378 patients (37.8% response rate) and 267 referring orthopedists (16.2% response rate) participated in our study and returned the survey. The following analysis reports on the 322 patients and 187 orthopedists who fully completed the DCE component of the questionnaire and provided consistent responses (see Table 2 for the demographic characteristics of both study groups). The mean age of patients was 68.99 (SD 10.44) years. Slightly more than half of patients (51.6%) were female, and most patients (53.7%) stated secondary general school or less as the highest educational level. Approximately one in four patients was aware of HRCs when answering the survey (26.1%). The mean age of participating orthopedists was 53.60 (SD 8.14) years, a large majority of respondents were male (90.4%), and 39 respondents (21.2%) indicated prior HRC experience in searching for a hospital. Almost three out of four surveyed referring orthopedists (72.7%) stated that they perceived substantial differences in the quality of care between hospitals.

Descriptive rating and ranking results

The results regarding the importance of different information items for the hospital choice—on a 1–5 scale (1 = not all important; 5 = extremely important)—showed that patients rated confirmed diagnosis (hip surgery) rate (4.65±0.80), mobility at hospital discharge (4.60±0.73), and complication rate (4.54±0.81) as most important (see Supplemental material 2). In contrast, referring physicians rated complication rate (4.19±1.06), the number of cases treated (4.09±0.98), and the 1-year revision surgery rate (4.00±1.21) as most important. Furthermore, the survey results for the single most important information item for the hospital choice revealed confirmed diagnosis (hip surgery) rate (44.5%) as well as the number of cases treated (35.9%) as most relevant from the perspective of patients and referring physicians, respectively.

Table 2 Key characteristics for both patients ($n = 322$) and referring physicians ($n = 187$)

	Patients ($n = 322$)		Referring physicians ($n = 187$)	
	Mean (SD) or n	Range or %	Mean (SD) or n	Range or %
Age				
Mean (SD)	68.99 (10.44)	35–89	53.60 (8.14)	33–73
Gender				
Male	151	46.9%	169	90.4%
Female	166	51.6%	18	9.6%
Missing	5	1.6%	0	0.0%
Educational attainment				
Secondary general school or less	173	53.7%	n.a	n.a
Intermediate secondary school	93	28.9%	n.a	n.a
(Technical) University entrance qualification	38	11.8%	n.a	n.a
Missing	18	5.6%	n.a	n.a
Chronic conditions				
Any chronic condition	141	43.8%	n.a	n.a
No chronic condition	159	49.4%	n.a	n.a
Missing	22	6.8%	n.a	n.a
Health Status				
Good or better	168	52.2%	124	66.3%
Satisfactory	122	37.9%	51	27.3%
Bad or worse	26	8.1%	12	6.4%
Missing	6	1.9%	0	0.0%
Perceived differences in the hospital quality				
Big differences	81	25.2%	136	72.7%
Small differences	89	27.6%	48	25.7%
No differences	23	7.1%	2	1.1%
I don`t know	122	37.9%	1	0.5%
Missing	7	2.2%	0	0.0%
Knowledge of hospital report cards				
Yes	84	26.1%	39	21.2%
No	188	56.4%	145	78.8%
Missing	50	15.5%	0	0.0%
Practice type				
Single physician practice	n.a	n.a	65	34.8%
Group physician practice	n.a	n.a	76	40.6%
Outpatient-based health-care center	n.a	n.a	40	21.4%
Other	n.a	n.a	6	3.2%
Attending doctor (i.e., with hospital affiliation)				
Yes and I perform the surgery myself	n.a	n.a	40	21.4%
Yes but I do not perform the surgery myself	n.a	n.a	21	11.2%
No	n.a	n.a	124	66.3%
Missing	n.a	n.a	2	1.1%

Results of the discrete choice experiments

The applied models displayed significant coefficients for all attributes included in both analyses ($p < 0.001$ each). This means that, independent of the placement, all attributes were relevant to the decision of both patients and referring physicians (Supplemental material 3). [Please also note that some mean levels could not been shown to be statistically significant; for example, the level "No assessment intended/results not (yet) available" did not reach statistical significance in case of the "Confirmed diagnosis rate" and the "Mobility of hospital discharge".] Among patients, the highest preference weights for hospital choice were calculated for "Postoperative complication rate" (coef.: 0.560; SE: 0.037), ''Mobility

at hospital discharge" (coef.: 0.551; SE: 0.036), and ''The number of cases treated" (coef.: 0.547; SE: 0.036). The fourth position was occupied by ''Confirmed diagnosis (hip surgery) rate'' (Coef.: 0.437; SE: 0.033), followed by ''Prevention of falls measures'' (Coef. 0.431; SE: 0.035), and ''Endocert Certificate'' (Coef. 0.341; SE: 0.033). In contrast, referring physicians showed a clear preference for a decrease in ''One-year revision surgery rate'' (coef.: 1.014; SE: 0.053). The following positions were occupied by "The number of cases treated" (coef.: 0.696; SE: 0.051) and "Postoperative complication rate" (coef.: 0.546; SE: 0.050). Here, the "Confirmed diagnosis (hip surgery) rate" (coef.: 0.390; SE: 0.052), "Endocert Certificate'' (coef.: 0.310; SE: 0.047), and the ''Mobility at hospital discharge'' (coef.: 0.236; SE: 0.046) were less important for referring physicians. The same order can be derived from the calculation of the relative importance of all attributes for both study groups. The relative importance was calculated based on each attribute's coefficient range (i.e., the difference between the coefficients of the highest and lowest level of each attribute) expressed as a share of the total range across attributes (Table 3). Table 3 also presents the odds ratio (OR) for each level to analyze its impact on the hospital choice. For example, the likelihood of voting a hospital with a successful result regarding the "Confirmed diagnosis rate" (i.e., quality targets reached) is 1.535 times higher than the probability of choosing a hospital with the mean of "confirmed diagnosis rate".

As shown in Fig. 2, the preference pattern of patients showed a relatively small range of results. The following three attributes, "Postoperative complication rate" (20.6%; level range of 1.164), "Mobility at hospital discharge" (19.9%; level range of 1.127), and "The number of cases treated" (18.5%; level range of 1.045) were weighted highest. The lowest weight was detected for "Endocert Certificate" (11.4%; level range of 0.646). In contrast, the preference pattern of referring orthopedists showed larger differences in the relative importance of presented quality information measures. Here, the "One-year revision surgery rate" was identified as the main dominant attribute (30.4%; level range of 1.989). Furthermore, "The number of cases treated" (21.0%; level range of 1.372) and "Postoperative complication rate" (17.2%; level range of 1.123) were relatively highly weighted. Here, the lowest weight was detected for "Mobility at hospital discharge" (8.2%; level range of 0.536).

Discussion

The objective of the present study was to identify the most important publicly reported hospital quality information for hospital choice and to determine their relative importance among both former patients and referring physicians

in Germany. The results might assist us in learning how to weight different quality measures when calculating aggregated patient-centered composite measures.

Based on our results, several conclusions can be drawn. First, patients and referring physicians selected similar publicly available quality information for the hospital choice. However, it should be mentioned that we focused on publicly available quality information and did not include information which is not available for public reporting purposes or which is not routinely collected in the health-care system in our experiments. Therefore, we cannot exclude that the integration of such information might have led to different findings. For example, Geraedts and colleagues showed that many surgical patients decided in favor of a hospital because of the trust they had built up through previous treatment in the hospital, which they experienced as satisfactory [80].

Second, HRCs should present the information that consumers value most when making hospital choices. At least five of the seven most relevant publicly reported quality information items are displayed on most German HRCs (e.g., Weisse Liste, AOK Hospital Navigator) [28]. This is mainly due to the fact that those information items are part of the German external hospital quality assurance system whose results are made available for public reporting purposes [81]. In contrast, the certification as a Certified Endoprosthetics Center is only presented on a few HRCs so far (e.g., Weisse Liste). Besides, the 1-year revision rate is only presented on the hospital navigator of the Allgemeine Ortskrankenkasse (AOK) [81]. The AOK is a major provider of statutory health insurance in the German health-care system and calculates further quality information based on claims data [29]. Following this, providers of HRCs might compare whether the most relevant quality information is already included on its own HRC. If not, providers might consider adjusting the content as a first step toward more consumer orientation.

Third, the study's DCE-based findings appear broadly consistent with the results from the rating-based survey. For example, the preference pattern of patients showed a relatively small range of results. While the rating-based findings (on a scale of 1–5) varied between 4.09 ± 1.03 (Endocert Certificate) and 4.65 ± 0.80 (confirmed diagnosis hip surgery rate), the DCE-based relative importance values ranged between 11.4% ("Endocert Certificate") and 20.6% (Postoperative complication rate), respectively. In both cases, "Endocert Certificate" was the least important attribute for the hospital choice while "Complication rate" and "Mobility at hospital discharge" were among the most important attributes for the hospital choice. However, the ranking-based results showed slightly different findings. While the confirmed diagnosis (hip surgery) rate (44.5%) was shown to be the single most important information item for the hospital choice, the DCE-based

**Table 3** Estimated parameters of multinomial logit (MNL) models

Attributes and levels	Patients (n = 322)							Referring physicians (n = 187)						
	Coeff	OR	SE coeff	95% CI		p	Difference (highest—lowest level)	Coeff	OR	SE coeff	95% CI		p	Difference (highest—lowest level)
Confirmed diagnosis rate[§,#]														
Quality targets reached	0.429	1.535	0.040	0.350	0.507	<.001	0.856	0.454	1.575	0.058	0.340	0.568	<.001	0.800
No assessment intended/results not (yet) available	−0.001	0.999	0.039	−0.078	0.076	0.977		−0.108	0.897	0.056	−0.218	0.001	0.051	
Quality targets not reached	−0.427	0.652	0.038	−0.503	−0.352	<.001		−0.346	0.708	0.062	−0.467	−0.224	<.001	
#Endocert Certificate[§]														
No certificate	−0.390	0.677	0.040	−0.469	−0.311	<.001	0.646	−0.446	0.640	0.058	−0.561	−0.331	<.001	0.718
Certified EndoProstheticsCenter	0.135	1.144	0.040	0.055	0.214	0.001		0.174	1.191	0.054	0.068	0.281	0.001	
Certified EndoProstheticsCenter of Maximum Care	0.256	1.291	0.039	0.179	0.332	<.001		0.272	1.312	0.059	0.155	0.388	<.001	
The number of cases treated[§,#]														
84 patients (below average)	−0.616	0.540	0.040	−0.695	−0.537	<.001	1.045	−0.676	0.509	0.058	−0.791	−0.560	<.001	1.372
148 patients (average)	0.188	1.207	0.039	0.112	0.264	<.001		−0.020	0.980	0.056	−0.131	0.090	0.716	
230 patients (above average)	0.428	1.535	0.044	0.341	0.516	<.001		0.696	2.006	0.058	0.581	0.811	<.001	
Postoperative complication rate[§,#]														
Quality targets reached	0.622	1.863	0.042	0.540	0.704	<.001	1.164	0.600	1.823	0.061	0.481	0.720	<.001	1.123
No assessment intended/results not (yet) available	−0.080	0.923	0.039	−0.157	−0.003	0.040		−0.078	0.925	0.057	−0.190	0.034	0.170	
Quality targets not reached	−0.542	0.582	0.044	−0.628	−0.455	<.001		−0.522	0.593	0.057	−0.636	−0.409	<.001	
Mobility at hospital discharge[§,#]														
Quality targets reached	0.597	1.817	0.043	0.513	0.682	<.001	1.127	0.344	1.410	0.057	0.231	0.456	<.001	0.536
No assessment intended/results not (yet) available	−0.068	0.935	0.039	−0.145	0.010	0.085		−0.192	0.826	0.061	−0.313	−0.071	0.002	
Quality targets not reached	−0.530	0.589	0.041	−0.610	−0.449	<.001		−0.152	0.859	0.056	−0.263	−0.042	0.007	
Prevention of falls measures[§]														
Quality targets reached	0.458	1.581	0.039	0.381	0.535	<.001	0.822	n.a	n.a	n.a	n.a	n.a	n.a	n.a
No assessment intended/results not (yet) available	−0.094	0.911	0.040	−0.172	−0.016	0.018		n.a	n.a	n.a	n.a	n.a	n.a	
Quality targets not reached	−0.364	0.695	0.042	−0.447	−0.281	<.001		n.a	n.a	n.a	n.a	n.a	n.a	
1-year revision surgery rate[#]														
1.41% (better than average)	n.a	n.a	n.a	n.a	n.a	n.a	n.a	0.977	2.657	0.094	0.791	1.163	<.001	1.989
2.48% (average)	n.a	n.a	n.a	n.a	n.a	n.a		0.035	1.035	0.159	−0.279	0.348	0.827	
3.89% (lower than average)	n.a	n.a	n.a	n.a	n.a	n.a		−1.012	0.363	0.097	−1.203	−0.821	<.001	
Model constant (Intercept)	−0.309	0.734	0.047	−0.402	−0.216	<.001		n.a	n.a	n.a	n.a	n.a	n.a	

OR odds ratio (OR = exp(coeff)); SE standard errors, CI confidence interval

[§]This attribute applies for patients

[#]This attribute applies for referring physicians

Fig. 2 Graphic display of level estimates with 95% CI (discrete choice experiment) and mean relative importance of attributes for both subgroups for the hospital choice
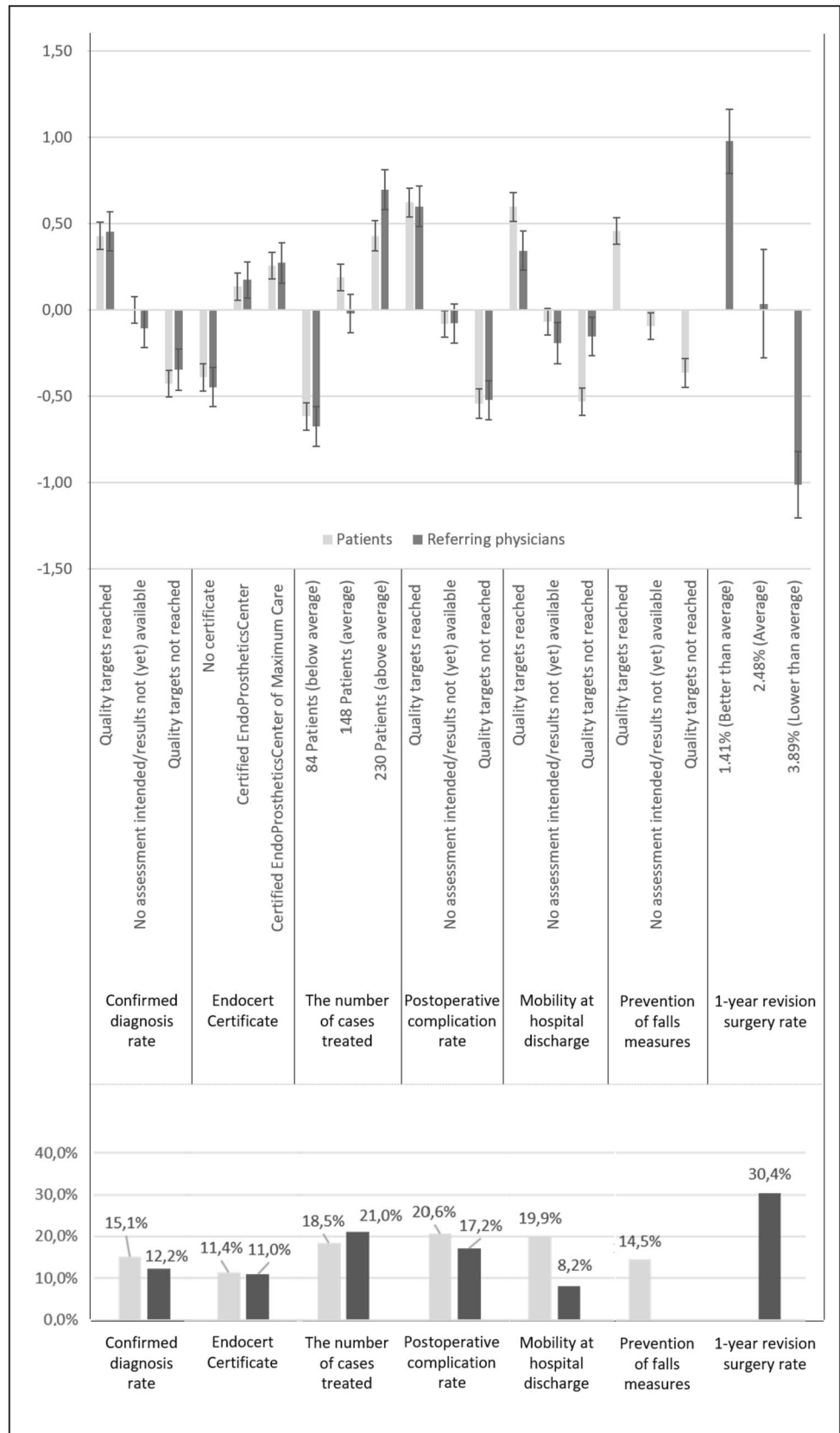
findings revealed a relative importance value of 15.1%. In addition, the overall results regarding the preference pattern of referring physicians were shown to be very consistent, regardless of the methodology. For example, the three attributes "The number of cases treated", "Complication rate", and "1-year revision surgery rate" showed the strongest results. In contrast, the "Confirmed diagnosis (hip surgery) rate", "Endocert Certificate", and "Mobility at hospital discharge" were shown to be less important.

Fourth, the results of this study could be used for the calculation of composite measures [11, 13–17]. Therefore, we applied a consumer-based weighting approach as recommended and determined the relative weight of relevant information [16]. For example, the preference pattern of referring orthopedists showed the "One-year revision surgery rate" to be most important with a relative weight of 30.4%. The challenge now is to convert the hospital quality results into (numeric) scores for calculating the composite measure. As stated by van Til and colleagues [82], the overall preference for alternatives (i.e., hospitals) might be estimated based on the sum of the part-worth utilities for the selected level of all attributes. For example, a clinic certified as an Endoprosthetics Center of Maximum Care (coef.: 0.272), with a below-average 1-year revision surgery rate (coef.: 0.977), an above-average number of cases (coef.: 0.696), that also achieves the statutory quality targets regarding confirmed diagnosis rate (coef.: 0.454), postoperative complication rates (coef.: 0.600), and mobility at hospital discharge (coef.: 0.344), would have an overall score of 3.343 from the perspective of referring physicians. In contrast, an overall score of -3.194 would be calculated for a hospital that achieves the most negative level expressions for each attribute. [Please consider this example to be simplified. A more advanced approach would consider the corresponding utility function to estimate the expected overall utility of each alternative (i.e., hospital).] Based on this, hospitals could be ranked and then grouped into several performance groups. Therefore, our study has shown that consumer preferences might differ substantially between target groups and each preference structure should be considered individually when developing rating systems based on composite measures. However, it should also be noted that our approach has certain limitations that must be considered carefully. For example, the attribute levels are used to operationalize the alternatives included in the choice sets. Therefore, it is important to select attributes and corresponding levels that properly describe the health-care product or service (i.e., the hospitals). If the levels are not defined in the appropriate range, the estimated coefficients could be biased [83]. Other methodological issues might also have an impact on the results derived by DCE studies and should be considered carefully [83, 84].

Finally, the results might serve as the basis for improved physician–patient communication. The findings might increase the mutual understanding of hospital choice preferences. In particular, referring physicians might better understand what patients actually want and which hospital they would prefer. They might have an important "agent" function in the hospital choice especially for patients with little time prior to admission and those who do not decide themselves [85]. Therefore, by understanding factors that influence patients' hospital choice decisions, referring physicians may become more sensitive to patients' preferences, which may have a positive effect on the process and outcomes of shared decision making [86]. This would increase the likelihood that patients may be satisfied with both the hospital choice process and the hospital choice itself [87]. Our study makes some contributions to such understanding, since it shows that, in contrast to referring physicians, the 1-year revision rate did not influence patients' stated preferences, whereas the prevention of falls measures did.

As with any study, there are several limitations that have to be considered. First, it is important to mention that this study was conducted in Germany and might be of limited relevance for other countries. Nevertheless, the results presented in this paper are of interest for all countries with public reporting initiatives, such as the United States and others. For example, Hospital Compare in the United States already provides an overall composite measure for hospitals but does not consider preferences of patients and referring physicians in detail [20]. Second, it must be noted that the selected attributes for the DCEs referred to publicly available quality measures. Therefore, we cannot exclude that other quality measures are more relevant from the perspective of both study groups. Nevertheless, it seems reasonable for our study purpose to follow the approach used. In comparable studies, very similar approaches have been applied [47]. Third, our study analyses stated that preferences of the respondents to hypothetical scenarios and their actual responses may differ [47]. Fourth, the sample of surveyed referring physicians ($n = 187$) was smaller than that of patients surveyed ($n = 322$). In this regard, it should be noted that we aimed to reach a minimum number of 150 participants in both studies to meet Johnson's often-used rule-of-thumb mentioned above [73], taking into account more advanced recommendations for statistical robustness [74, 75]. Following this, we were able to achieve our set target regarding the recruitment in both studies. Finally, the comprehensibility of the study attributes and corresponding levels is crucial for conducting DCEs and analyzing the corresponding data. During our study, we aimed to address this topic by means of different approaches. First, we qualitatively interviewed 20 patients who had undergone elective hip arthroplasty surgery within 3 years

prior to answering the survey. Thereby, we evaluated the comprehensibility of the labels used for the attributes and their corresponding descriptions in particular to assure the intended interpretation of the explanations. Based on those interviews, slight modifications were made to eliminate remaining uncertainties. Second, the respondents had to work with the presented study attributes and corresponding levels before conducting the choice-based part of the DCE. Both approaches might have positively influenced the comprehensibility of the study attributes and levels. In addition, the surveyed patient sample consisted of patients who had undergone elective hip arthroplasty surgery within 3 years prior to answering the survey. This might indicate—at least to a certain amount—some experience with hip-surgery-related issues.

In sum, this study adds to the literature by presenting results of 2 surveys among 322 patients and 187 referring physicians from the German outpatient sector to identify the most important publicly reported hospital quality information for hospital choice and to determine their relative importance. The results might support us in learning how to weight different quality measures when calculating aggregated composite measures based on consumer-based weighting. It could be shown that patients and referring physicians have similar interests with regard to the selection of publicly available quality information for the choice of hospitals. Nevertheless, we saw meaningful differences between patients and referring physicians in the relative weight of relevant information. This might have a significant impact on the hospital ranking results for each study group. Future research should address in detail how to convert hospital quality results into (numeric) scores for calculating composite measures. Besides, by understanding factors that influence patients' hospital choice decisions, referring physicians may become more sensitive to patients' preferences, which may have a positive effect on the process and outcomes of shared decision making.

Declarations

References

1. Totten, A.M., Wagner, J., Tiwari, A., O`Haire, C., Griffin, J., Walker, M.: Public reporting as a quality improvement strategy: closing the quality gap: revisiting the state of the science. In: AHRQ Publication No. 12-E011-EF. Evidence Report/Technology Assessment; Number 208. (2012)

2. Berwick, D.M., James, B., Coye, M.J.: Connections between quality measurement and improvement. Med. Care **41**(1), I30–I38 (2003)

3. Contandriopoulos, D., Champagne, F., Denis, J.-L.: The multiple causal pathways between performance measures' use and effects. Med. Care Res. Rev. **71**(1), 3–20 (2014). https://doi.org/10.1177/1077558713496320

4. Emmert, M., Schindler, A., Drach, C., et al.: The use intention of hospital report cards among patients in the presence or absence of patient-reported outcomes. Health Policy **126**(6), 541–548 (2022). https://doi.org/10.1016/j.healthpol.2022.03.015

5. Hussey, P.S., Luft, H.S., McNamara, P.: Public reporting of provider performance at a crossroads in the United States: summary of current barriers and recommendations on how to move forward. Med. Care Res. Re. MCRR **71**(5 Suppl), 5S-16S (2014). https://doi.org/10.1177/1077558714535980

6. Metcalfe, D., Rios Diaz, A.J., Olufajo, O.A., et al.: Impact of public release of performance data on the behaviour of healthcare consumers and providers. Cochrane Database Syst. Rev. **9**, CD004538 (2018). https://doi.org/10.1002/14651858.CD004538.pub3

7. Haviland, A.M., Damberg, C.L., Mathews, M., Paddock, S.M., Elliott, M.N.: Shifting from passive quality reporting to active

nudging to influence consumer choice of health plan. Med. Care Res. Rev. (2018). https://doi.org/10.1177/1077558718798534

8. de Groot, I.B., Otten, W., Dijs-Elsinga, J., Smeets, H.J., Kievit, J., Marang-van de Mheen, P.J.: Choosing between hospitals: the influence of the experiences of other patients. Med. Decis. Making 32(6), 764–778 (2012). https://doi.org/10.1177/0272989X12443416

9. Pross, C., Busse, R., Geissler, A.: Hospital quality variation matters—a time-trend and cross-section analysis of outcomes in German hospitals from 2006 to 2014. Health Policy 121(8), 842–852 (2017). https://doi.org/10.1016/j.healthpol.2017.06.009

10. Strumann, C., Geissler, A., Busse, R., Pross, C.: Can competition improve hospital quality of care? A difference-in-differences approach to evaluate the effect of increasing quality transparency on hospital quality. Eur. J. Health Econ. 23(7), 1229–1242 (2022). https://doi.org/10.1007/s10198-021-01423-9

11. Sandmeyer, B., Fraser, I.: New evidence on what works in effective public reporting. Health Serv. Res. 51(Suppl 2), 1159–1166 (2016). https://doi.org/10.1111/1475-6773.12502

12. Schlesinger, M., Kanouse, D.E., Martino, S.C., Shaller, D., Rybowski, L.: Complexity, public reporting, and choice of doctors: a look inside the blackest box of consumer behavior. Med. Care Res. Rev. 71(5 Suppl), 38S-64S (2014). https://doi.org/10.1177/1077558713496321

13. Smith, M.A., Bednarz, L., Nordby, P.A., et al.: Increasing consumer engagement by tailoring a public reporting website on the quality of diabetes care: a qualitative Study. JMIR. 18(12), e332 (2016). https://doi.org/10.2196/jmir.6555

14. Huckman, R.S., Kelley, M.A.: Public reporting, consumerism, and patient empowerment. N. Engl. J. Med. 369(20), 1875–1877 (2013). https://doi.org/10.1056/NEJMp1310419

15. Sinaiko, A.D., Eastman, D., Rosenthal, M.B.: How report cards on physicians, physician groups, and hospitals can have greater impact on consumer choices. Health Affairs (Project Hope). 31(3), 602–611 (2012). https://doi.org/10.1377/hlthaff.2011.1197

16. Shwartz, M., Restuccia, J.D., Rosen, A.K.: Composite measures of health care provider performance: a description of approaches. Milbank Q. 93(4), 788–825 (2015). https://doi.org/10.1111/1468-0009.12165

17. Shwartz, M., Rosen, A.K., Burgess, J.F.: Can composite measures provide a different perspective on provider performance than individual measures? Med. Care 55(12), e99–e103 (2017). https://doi.org/10.1097/MLR.0000000000000407

18. Werner, R.M., Konetzka, R.T., Polsky, D.: Changes in consumer demand following public reporting of summary quality ratings: an evaluation in nursing homes. Health Serv. Res. 51(Suppl 2), 1291–1309 (2016). https://doi.org/10.1111/1475-6773.12459

19. Bhandari, N., Scanlon, D.P., Shi, Y., Smith, R.A.: Why do so few consumers use health care quality report cards? A framework for understanding the limited consumer impact of comparative quality information. Med. Care Res. Rev. 76(5), 515–537 (2019). https://doi.org/10.1177/1077558718774945

20. The Centers for Medicare and Medicaid Services: Overall hospital quality star rating. https://data.cms.gov/provider-data/topics/hospitals/overall-hospital-quality-star-rating/#measure-included-by-categories.

21. Hibbard, J.H., Greene, J., Daniel, D.: What is quality anyway? Performance reports that clearly communicate to consumers the meaning of quality of care. Med. Care Res. Rev. 67(3), 275–293 (2010). https://doi.org/10.1177/1077558709356300

22. Emmert, M., Meier, F., Pisch, F., Sander, U.: Physician choice making and characteristics associated with using physician-rating websites: cross-sectional study. J. Med. Internet Res. 15(8), e187 (2013). https://doi.org/10.2196/jmir.2702

23. Mukamel, D.B., Weimer, D.L., Zwanziger, J., Gorthy, S.-F.H., Mushlin, A.I.: Quality report cards, selection of cardiac surgeons, and racial disparities: a study of the publication of the New York State Cardiac Surgery Reports. Inquiry 41(4):435–446 (2004–2005)

24. Emmert, M., Meszmer, N., Jablonski, L., Zinth, L., Schöffski, O., Taheri-Zadeh, F.: Public release of hospital quality data for referral practices in Germany: results from a cluster-randomised controlled trial. Health Econ Rev. 7(33), I30 (2017). https://doi.org/10.1186/s13561-017-0171-5

25. Geraedts, M., Hermeling, P., Ortwein, A., de Cruppé, W.: Public reporting of hospital quality data: what do referring physicians want to know? Health Policy 122(11), 1177–1182 (2018). https://doi.org/10.1016/j.healthpol.2018.09.010

26. Salampessy, B.H., Ikkersheim, D., Portrait, F.R.M., Koolman, X.: Do patients' preferences prevail in hospital selection? A comparison between discrete choice experiments and revealed hospital choice. BMC Health Serv. Res. 22(1), 1136 (2022). https://doi.org/10.1186/s12913-022-08403-6

27. Berwick, D.M., James, B., Coye, M.J.: Connections between quality measurement and improvement. Med. Care 41(1 Suppl), I30–I38 (2003)

28. Emmert, M., Meszmer, N., Simon, A., Sander, U.: Internetportale für die Krankenhauswahl in Deutschland: Eine leistungsbereichsspezifische Betrachtung. Gesundheitswesen (2015). https://doi.org/10.1055/s-0035-1549968

29. Wissenschaftliches Instituts der AOK: Qualitätssicherung mit Routinedaten. https://www.qualitaetssicherung-mit-routinedaten.de/.

30. Institut für Qualitätssicherung und Transparenz im Gesundheitswesen: Hüftendoprothesenversorgung - Bundesauswertung zum Erfassungsjahr 2020: Qualitätsindikatoren und Kennzahlen. Berlin (2021). https://iqtig.org/downloads/auswertung/2020/hep/QSKH_HEP_2020_BUAW_V01_2021-08-10.pdf.

31. Determann, D., Gyrd-Hansen, D., de Wit, G.A., et al.: Designing unforced choice experiments to inform health care decision making: implications of using opt-out, neither, or status quo alternatives in discrete choice experiments. Med. Decis. Making 39(6), 681–692 (2019). https://doi.org/10.1177/0272989X19862275

32. Vass, C., Rigby, D., Payne, K.: The role of qualitative research methods in discrete choice experiments. Med. Decis. Making 37(3), 298–313 (2017). https://doi.org/10.1177/0272989X16683934

33. McFadden, D.: Conditional logit analysis of qualitative choice behavior. In: Frontiers in econometrics, pp 105–142. Academic Press, New York (1974)

34. Lancaster, K.J.: A new approach to consumer theory. J. Polit. Econ. LXXIV, 2 (1966)

35. Naik Panvelkar, P., Armour, C., Saini, B.: Community pharmacy-based asthma services—what do patients prefer? J. Asthma 47(10), 1085–1093 (2010). https://doi.org/10.3109/02770903.2010.514638

36. Reed Johnson, F., Lancsar, E., Marshall, D., et al.: Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. Value Health. 16(1), 3–13 (2013). https://doi.org/10.1016/j.jval.2012.08.2223

37. Bridges, J.F.P., Hauber, A.B., Marshall, D., et al.: Conjoint analysis applications in health–a checklist: a report of the ISPOR good research practices for conjoint analysis task force. Value Health. 14(4), 403–413 (2011). https://doi.org/10.1016/j.jval.2010.11.013

38. Hauber, A.B., González, J.M., Groothuis-Oudshoorn, C.G.M., et al.: Statistical methods for the analysis of discrete choice experiments: a report of the ISPOR conjoint analysis good research

practices task force. Value Health. **19**(4), 300–315 (2016). https://doi.org/10.1016/j.jval.2016.04.004

39. Higgins, J., Thomas, J., Chandler, J., et al. (eds): Cochrane Handbook for Systematic Reviews of Interventions. 2nd Edition. In: Cochrane Book Series. John, Chichester (2019)

40. Ketelaar, N.A.B.M., Faber, M.J., Braspenning, J.C., Westert, G.P.: Patients' expectations of variation in quality of care relates to their search for comparative performance information. BMC Health Serv. Res. **14**, 617 (2014). https://doi.org/10.1186/s12913-014-0617-y

41. Zwijnenberg, N., Damman, O., Spreeuwenberg, P., Hendriks, M., Rademakers, J.: Different patient subgroup, different ranking? Which quality indicators do patients find important when choosing a hospital for hip- or knee arthroplasty? BMC Health Serv. Res. **11**, 299 (2011). https://doi.org/10.1186/1472-6963-11-299

42. Blake, R.S., Clarke, H.D.: Hospital compare and hospital choice: public reporting and hospital choice by hip replacement patients in texas. Med. Care Res. Rev. **76**(2), 184–207 (2019). https://doi.org/10.1177/1077558717699311

43. Beukers Puck, D.C., Kemp Ron, G.M., Varkevisser, M.: Patient hospital choice for hip replacement: empirical evidence from the Netherlands. Eur. J. Health Econ. HEPAC Health Econ. Prevent. Care **15**(9), 927–936 (2014). https://doi.org/10.1007/s10198-013-0535-7

44. Gutacker, N., Siciliani, L., Moscelli, G., Gravelle, H.: Choice of hospital: which type of quality matters? J. Health Econ. **50**, 230–246 (2016). https://doi.org/10.1016/j.jhealeco.2016.08.001

45. Moscelli, G., Siciliani, L., Gutacker, N., Gravelle, H.: Location, quality and choice of hospital: evidence from England 2002–2013. Reg. Sci. Urban Econ. **60**, 112–124 (2016). https://doi.org/10.1016/j.regsciurbeco.2016.07.001

46. Damman, O.C., Spreeuwenberg, P., Rademakers, J., Hendriks, M.: Creating compact comparative health care information: what are the key quality attributes to present for cataract and total hip or knee replacement surgery? Med. Decis. Making **32**(2), 287–300 (2012). https://doi.org/10.1177/0272989X11415115

47. O'Hara, L.M., Caturegli, I., O'Hara, N.N., et al.: What publicly available quality metrics do hip and knee arthroplasty patients care about most when selecting a hospital in Maryland: a discrete choice experiment. BMJ Open **9**(5), e028202 (2019). https://doi.org/10.1136/bmjopen-2018-028202

48. Moser, A., Korstjens, I., van der Weijden, T., Tange, H.: Patient's decision making in selecting a hospital for elective orthopaedic surgery. J. Eval. Clin. Pract. **16**(6), 1262–1268 (2010). https://doi.org/10.1111/j.1365-2753.2009.01311.x

49. Kandrack, R., Mehrotra, A., DeVries, A., Wu, S.-J., SooHoo, N.F., Martsolf, G.R.: Patient use of cost and quality data when choosing a joint replacement provider in the context of reference pricing. Health Serv. Res. Manag. Epidemiol. (2015). https://doi.org/10.1177/2333392815598310

50. Kolb, B., Emmert, M., Sander, U., Patzelt, C., Schöffski, O.: Do German public reporting websites provide information that office-based physicians consider before referring patients to hospital? A four-step analysis. Z. Evid. Fortbild. Qual. Gesundhwes. (2018). https://doi.org/10.1016/j.zefq.2018.07.010

51. Ikkersheim, D., Koolman, X.: The use of quality information by general practitioners: does it alter choices? A randomized clustered study. BMC Fam. Pract. **14**, 95 (2013). https://doi.org/10.1186/1471-2296-14-95

52. Ringard, A.: Why do general practitioners abandon the local hospital? An analysis of referral decisions related to elective treatment. Scand. J. Public Health. **38**(6), 597–604 (2010). https://doi.org/10.1177/1403494810371019

53. Prang, K.-H., Canaway, R., Bismark, M., Dunt, D., Kelaher, M.: The use of public performance reporting by general practitioners: a study of perceptions and referral behaviours. BMC Fam. Pract. **19**(1), 29 (2018). https://doi.org/10.1186/s12875-018-0719-4

54. Rotar, A.M., van den Berg, M.J., Schäfer, W., Kringos, D.S., Klazinga, N.S.: Shared decision making between patient and GP about referrals from primary care: Does gatekeeping make a difference? PLoS ONE **13**(6), e0198729 (2018). https://doi.org/10.1371/journal.pone.0198729

55. Potappel, A.J.C., Meijers, M.C., Kloek, C., et al.: To what degree do patients actively choose their healthcare provider at the point of referral by their GP? A video observation study. BMC Fam. Pract. **20**(1), 166 (2019). https://doi.org/10.1186/s12875-019-1060-2

56. Beckert, W.: Choice in the presence of experts: the role of general practitioners in patients' hospital choice. J. Health Econ. **60**, 98–117 (2018). https://doi.org/10.1016/j.jhealeco.2018.06.003

57. Birk, H.O., Henriksen, L.O.: Which factors decided general practitioners' choice of hospital on behalf of their patients in an area with free choice of public hospital? A questionnaire study. BMC Health Serv. Res. **12**, 126 (2012). https://doi.org/10.1186/1472-6963-12-126

58. Doering, N., Maarse, H.: The use of publicly available quality information when choosing a hospital or health-care provider: the role of the GP. Health Expect. **18**(6), 2174–2182 (2015). https://doi.org/10.1111/hex.12187

59. Geraedts, M., Schwartze, D., Molzahn, T.: Hospital quality reports in Germany: patient and physician opinion of the reported quality indicators. BMC Health Serv. Res. **7**, 157 (2007). https://doi.org/10.1186/1472-6963-7-157

60. Frezza, E.E., Wachtel, M.S.: The quality of a surgeon defined by internal medicine and family practice physicians: a closed-ended survey with importance scale. Am. Surg. **73**(5), 481–483 (2007)

61. Ferrua, M., Sicotte, C., Lalloue, B., Minvielle, E.: Comparative quality indicators for hospital choice: do general practitioners care? PLoS ONE **11**(2), e0147296 (2016). https://doi.org/10.1371/journal.pone.0147296

62. Ketelaar, N.A., Faber, M.J., Elwyn, G., Westert, G.P., Braspenning, J.C.: Comparative performance information plays no role in the referral behaviour of GPs. BMC Fam. Pract. **15**, 146 (2014). https://doi.org/10.1186/1471-2296-15-146

63. Hermeling, P., de Cruppé, W., Geraedts, M.: Informationsbedarf niedergelassener Ärzte bei Ein- und Überweisungen. Gesundheitswesen **75**(7), 448–455 (2013). https://doi.org/10.1055/s-0032-1327745

64. Hesse, A., Schreyögg, J.: Determinanten eines erfolgreichen Einweisermarketings für Krankenhäuser - eine explorative Analyse. Gesundheitsokonomie und Qualitatsmanagement. **12**(5), 310–314 (2007). https://doi.org/10.1055/s-2007-962846

65. Konrad, S., Raab, A.: Welche Faktoren beeinflussen die Einweisungsentscheidung von niedergelassenen Ärzten in der Praxis? Gesundheitsokonomie und Qualitatsmanagement. **19**(03), 137–143 (2014). https://doi.org/10.1055/s-0034-1365968

66. Mühlbacher, A.C., Bethge, S.: Patients' preferences: a discrete-choice experiment for treatment of non-small-cell lung cancer. Eur. J. Health Econ. **16**(6), 657–670 (2015). https://doi.org/10.1007/s10198-014-0622-4

67. Haaijer, R., Wedel, M.: Conjoint choice experiments: general characteristics and alternative model specifications. In: Gustafsson. A. (ed.) Conjoint measurement: Methods and applications ; with 91 tables. 3. ed. Springer, Berlin, Heidelberg, New York (2003). http://digitale-objekte.hbz-nrw.de/webclient/DeliveryManager?pid=1487557&custom_att_2=simple_viewer

68. Cunningham, C.E., Deal, K., Chen, Y.: Adaptive choice-based conjoint analysis: a new patient-centered approach to the assessment of health service preferences. Patient. **3**(4), 257–273 (2010). https://doi.org/10.2165/11537870-000000000-00000

69. Dommaraju, S., Hagey, J., Odeny, T.A., et al.: Preferences of people living with HIV for differentiated care models in Kenya: a

discrete choice experiment. PLoS ONE **16**(8), e0255650 (2021). https://doi.org/10.1371/journal.pone.0255650

70. Ryan, M., Gerard, K.: Using discrete choice experiments to value health care programmes: current practice and future research reflections. Appl. Health Econ. Health Policy **2**(1), 55–64 (2003)

71. Sawtooth Software: The CBC system for choice-based conjoint analysis. In: Sawtooth Software, Inc., ed. CBC—Technical Paper [Technical Paper Series]. Washington D.C (2008)

72. The German Medical Association: Ärztestatistik zum 31. Dezember 2020. Berlin (2021). https://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/Statistik_2020/2020-Statistik.pdf.

73. Orme. B.K.: Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research. Fourth Edition (2020)

74. Orme. B.K., Chrzan, K.: Becoming an Expert in Conjoint Analysis: Choice Modeling for Pros. North Orem, UT (2017)

75. Mühlbacher, A., Bethge, S.: What matters in type 2 diabetes mellitus oral treatment? A discrete choice experiment to evaluate patient preferences. Eur. J. Health Econ. **17**(9), 1125–1140 (2016). https://doi.org/10.1007/s10198-015-0750-5

76. Croissant, Y.: Estimation of random utility models in R : the mlogit package. J. Stat. Soft. (2020). https://doi.org/10.18637/jss.v095.i11

77. Mühlbacher, A.C., Bridges, J.F.P., Bethge, S., et al.: Preferences for antiviral therapy of chronic hepatitis C: a discrete choice experiment. Eur. J. Health Econ. **18**(2), 155–165 (2017). https://doi.org/10.1007/s10198-016-0763-8

78. Sijstermans, E., Cheung, K.L., Goossens, A.J.M., Conde, R., Gonzalez, J.G.R., Hiligsmann, M.: A discrete choice experiment to assess patients' preferences for HIV treatment in the urban population in Colombia. J. Med. Econ. **23**(8), 812–818 (2020). https://doi.org/10.1080/13696998.2020.1735399

79. Eggers, F., Sattler, H., Teichert, T., Völckner, F.: Choice-based conjoint analysis. In: Homburg, C., Klarmann, M., Vomberg, A. (eds.) Handbook of market research, pp. 781–819. Springe, Cham (2022)

80. de Cruppé, W., Kummer, M.-C., Geraedts, M.: Wie wählen chirurgische patienten ihr krankenhaus? Eine querschnittstudie how do surgical patients choose their hospital? A cross-sectional study. Zentralbl. Chir. (2021). https://doi.org/10.1055/a-1555-3907

81. Emmert, M., Meszmer, N., Jablonski, L., Zinth, L., Schöffski, O., Taheri-Zadeh, F.: Public release of hospital quality data for referral practices in Germany: results from a cluster-randomised controlled trial. Health Econ. Rev. (2017). https://doi.org/10.1186/s13561-017-0171-5

82. van Til, J.A., Stiggelbout, A.M., Ijzerman, M.J.: The effect of information on preferences stated in a choice-based conjoint analysis. Patient Educ. Couns. **74**(2), 264–271 (2009). https://doi.org/10.1016/j.pec.2008.08.025

83. van den Broek-Altenburg, E., Atherly, A.: Using discrete choice experiments to measure preferences for hard to observe choice attributes to inform health policy decisions. Health Econ. Rev. (2020). https://doi.org/10.1186/s13561-020-00276-x

84. van der Pol, M., Currie, G., Kromm, S., Ryan, M.: Specification of the utility function in discrete choice experiments. Value Health. **17**(2), 297–301 (2014). https://doi.org/10.1016/j.jval.2013.11.009

85. de Cruppé, W., Geraedts, M.: Hospital choice in Germany from the patient's perspective: a cross-sectional study. BMC Health Serv. Res. **17**(1), 720 (2017). https://doi.org/10.1186/s12913-017-2712-3

86. de Bekker-Grob, E.W., Bliemer, M.C.J., Donkers, B., et al.: Patients' and urologists' preferences for prostate cancer treatment: a discrete choice experiment. Br. J. Cancer **109**(3), 633–640 (2013). https://doi.org/10.1038/bjc.2013.370

87. de Cruppé, W., Geraedts, M.: Wie wählen Patienten ein Krankenhaus für elektive operative Eingriffe? Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz **54**(8), 951–957 (2011). https://doi.org/10.1007/s00103-011-1320-3