

The Outcomes and Experiences Questionnaire: development and validation

Elizabeth Gibbons
Paul Hewitson
David Morley
Crispin Jenkinson
Ray Fitzpatrick

Health Service Research Unit, Nuffield
Department of Population Health,
University of Oxford, Oxford, UK

Background: This report presents evidence regarding the development and validation of a new questionnaire, the Outcomes and Experiences Questionnaire (OEQ). The rationale for the questionnaire is to bring together into one short instrument questions about two distinct domains – patients' reports of the outcomes of their care and how they experience care.

Methods: The OEQ was developed from literature reviews, iterative drafting and discussion within the research group and cognitive testing with a sample of patients who had a hospital experience. Two validation studies were carried out with an eleven item OEQ. The goals of the studies were to examine response rates and to test specific hypotheses of how OEQ should relate to other variables normally collected in the two studies. In the first study, the OEQ was added to the follow-up questionnaires for patients (n=490) receiving surgery for hip or knee replacement or varicose vein procedures participating in the national Patient Reported Outcome Measures (PROMs) program permitting the analysis of the OEQ against change scores for the measures obtained before and after surgery. In the second study the OEQ was included in a sample of patients (n=586) who had been selected to receive the National Health Service (NHS) inpatient survey from three contrasting hospital trusts.

Results: Results from study one provided consistent and substantial evidence of construct validity of OEQ particularly for those receiving hip or knee replacement. The OEQ sub-scales behaved differently and as predicted against other PROMs variables. Again hypotheses of how the two sub-scales regarding outcomes and experiences would relate to the existing domains of patient experience in the inpatient survey were broadly confirmed in study two.

Conclusion: The report provides encouraging evidence of the OEQ's capacity to assess distinct reports from patients about outcomes and experiences of care within a single short questionnaire.

Keywords: patient experience, patient reported outcome measures, NHS, questionnaire

Introduction

Monitoring how patients experience health care is a vital resource for improving services.^{1,2} It has also been increasingly accepted that, in addition to views about how they experience services, patients also have distinctive and important views about the outcomes of their health care. The need to capture these views of outcomes was identified as a priority in the future plans for the National Health Service (NHS) and a commitment was expressed to develop and roll out measures of patients' views of outcomes (Patient Reported Outcome Measures [PROMs]) wherever feasible.^{3,4} For the first time in the NHS, a national program was established mandating PROMs data collection for four very common elective procedures (hip and knee replacement; varicose vein [VV] surgery and groin surgery).

Correspondence: Elizabeth Gibbons
Health Service Research Unit, Nuffield
Department of Population Health,
University of Oxford, New Richards
Building, Roosevelt Drive, Old Road
Campus, Oxford, OX3 7LF, UK
Tel +44 1865 289 402
Email elizabeth.gibbons@dph.ox.ac.uk

Commonly, separate measures of these two distinct aspects of patients' responses have emerged. Questionnaires addressing how patients experience the care they receive, patient reported experience measures, have largely non-overlapping content with questionnaires to capture their views of the outcomes of that care. This can result in different arrangements to obtain, process, and feed back patients' views about these two aspects of their care. The likeliest result is that information about the two aspects of services is fed back through different mechanisms to different audiences requiring considerable effort to link views about different aspects of specific episodes of care.

From the patient's point of view, how he or she is cared for, and the outcomes of that care, are highly related and important aspects of a single episode or period of care. Similarly, as care is increasingly delivered in integrated and team-based services, it is difficult, from the perspective of the service, to justify separating out and creating different systems to capture feedback on experiences and outcomes in relation to episodes or periods of care.

The purpose of the study reported here was to develop and evaluate the psychometric properties of a questionnaire that would focus on these two aspects of patients' views about specific episodes or periods of their care: how care is experienced and its perceived outcomes.

Patients can provide feedback about an enormous array of different aspects of their care and health questionnaires have grown in length as well as diversity of forms to reflect the range of patients' experiences of health and health care. In scoping the ambition for the new questionnaire it was decided that it would need to address a small number of key issues that matter most to patients. There were several advantages to such a targeted approach. From the patient's point of view respondent burden should be minimized. A questionnaire that focused on key matters might be particularly acceptable to patients. Other evidence indicates the scope for shortening instruments without loss of important information and insight.⁵ In terms of use and impact on the service, it can be argued that attention from the health service will more readily be given to feedback on a small number of key aspects of services, particularly if there is reason to think they are priorities for the patient. It was envisaged that the questionnaire should be broadly applicable to patients admitted to hospital for treatment, rather than purely for investigation or assessment.

This article describes the process of developing the new questionnaire, termed "Outcomes and Experiences Questionnaire" (OEQ), and then reports the results of two studies to examine its construct validity.

Materials and methods

Design and development of the OEQ

The over-arching requirement for the new questionnaire was that a small number of items would address key aspects of outcomes and experiences when retrospectively assessing an episode or period of care. The goal of the project was therefore to provide a short questionnaire that addressed patients' perceptions of outcomes of an episode of care as well as how the care was experienced.

The development of "outcomes" (OEQ-O) items was informed by a literature review principally focusing on the conceptual and methodological issues related to patients' retrospective judgments of outcomes; details of the review have been published elsewhere, describing the range of retrospective judgments available such as comparisons of current with past health states.⁶ This included approaches such as patients' assessments and satisfaction with the outcomes of treatment as well as the application of transition items, anchor based methods, and global judgments of capturing those perceptions. The common strand in the items developed in relation to outcomes was the patient's ability to make retrospective judgments of outcomes of services.

The items comprising of the "experiences" domain (OEQ-E) drew upon available evidence of key dimensions of patient centered care: Institute of Medicine,⁷ and elaboration by the King's Fund Point of Care programme,⁸ and Picker Principles of Patient Centred Care.⁹ The purpose of the OEQ-E was to focus on those matters of concern to patients that Murrells et al describe as "relational" and captured aspects of care most closely related to clinical activity and interaction with health care professionals.¹⁰ Other domains of patient experience such as cleanliness, noise, and catering, although important, were more relevant to other audiences in management not principally targeted in OEQ. Items from existing patient experience questionnaires, principally the NHS patient surveys were reviewed. This included items that were at the time included in the 2010 Commissioning for Quality and Innovation payment framework for patient experience. The following domains were considered applicable to the relational aspects of care:

- communication;
- information;
- involvement;
- responsiveness to individual needs;
- discussion of worries and fears.

Items were identified from existing measures, assessed for relevance and re-drafted by experts within the research team.

Eleven items were agreed through iterative discussion by the group as reflecting a spectrum of distinct key aspects of patients' experiences of outcomes and care: outcomes (five items), experiences (six items), contributing to two sub-scales (Supplementary material). The OEQ-O is a summed scale adding the scores for the individual items: Q1, Q2, Q3, Q4, Q11. Scores range from 0 to 20 with a higher score indicative of a better outcome. The OEQ-E is a summed scale adding the scores for the individual items: Q5, Q6, Q7, Q8, Q9, and Q10. Scores range from 0 to 18 with high scores indicating a good experience.

Face and content validity

A panel of members of the public was established to provide feedback and gain views on the questionnaire evaluating face and content validity the extent to which the content of a scale is representative of the conceptual domain it is intended to cover; assessed qualitatively during the questionnaire development phase through pre-testing with patients.

Thirty people were recruited with diverse social and demographic backgrounds and a range of experiences of hospitalization in the previous year. Recruitment was from several sources including local members of the public recruited to support NHS/National Institute for Health Research (NIHR) and participants of research carried out by the Health Experiences Research Group (HERG) at Oxford. Adverts were distributed to other potential recruitment sources, for example, People in Research (<http://www.peopleinresearch.org/>). Snowball sampling identified other people to contact.

The 30 members of the public participated in cognitive testing of the eleven items of the OEQ. They had all had a hospital experience in the previous year. Three rounds of cognitive testing was conducted to test meaningfulness, clarity, and acceptability of successive versions of the test items of the OEQ. The eleven items were revised after each round of interviews. Round two included $n=15$ of the $n=30$ participants and round three, $n=8$ participants were interviewed.

Testing acceptability and construct validity

The next step was to examine response rates reflecting acceptability and aspects of construct validity of the new questionnaire. Evaluating construct validity provides evidence that the scale is correlated with other measures of the same or similar constructs in the hypothesized direction; assessed on the basis of correlations between the measure and other similar measures. High correlations between the

scale and relevant constructs preferably based on a priori hypothesis with predicted strength of correlation.

In particular it was important to examine the hypothesis that the OEQ measured two distinct aspects of patients' responses to services or whether, to the contrary, answers appeared to reflect more global and undifferentiated reactions to care. Two studies were designed to test the acceptability and construct validity of the newly created eleven items of the OEQ.

In the first study the performance of the novel items of the OEQ was to be examined when added to a sub-sample of patients participating in the NHS National PROMs program for elective surgery (hip and knee replacement and VV surgery). In the National PROMs program patients completed condition-specific health status instruments before and after surgery, for patients receiving hip replacement surgery, the Oxford Hip Score (OHS),¹¹ for patients receiving knee replacement surgery, the Oxford Knee Score (OKS),¹² and for patients receiving VV surgery, the Aberdeen Varicose Vein Questionnaire (AVVQ).¹³ Condition-specific instruments most directly addressed patients' views in relation to surgery received. For all three surgical procedures, patients also completed before and after surgery the EQ-5D, a generic measure addressing broader aspects of health.¹⁴ The National PROMs program also asked all respondents two single item questions (each with five response categories) at follow-up: how they would describe the results of the operation, range of responses: "excellent" to "poor" ("overall assessment") and how they viewed their problem at follow-up compared to before surgery, range of responses: "much better" to "much worse", ("perceived improvement").

A specific set of prior hypotheses were identified that allowed specification of expectations of various positive correlations between the new instrument and other variables. It was hypothesized that the OEQ-O would correlate more with the change scores for condition-specific instruments than with change scores for a generic health status measure, the EQ-5D, on the grounds that unlike the other two measures, the EQ-5D did not directly address the patient's episode specific views of outcomes. In other words patients' views about the benefits of (in this case) elective surgery should correlate most with changes in those aspects of their health status most immediately relevant to their surgery (condition-specific PROMs) than changes in a more general measure of health status.

In addition, it was hypothesized that the OEQ-E should correlate more weakly than would the OEQ-O with change scores for health status. The rationale behind these hypotheses

was that experiences of how services are delivered (eg, feeling that staff address personal concerns and give helpful information) are distinct aspects of care from the impact of treatment on outcomes and are distinguished in patients' responses. It was further hypothesized that the two single item questions from the National PROMs program (on overall assessment and perceived improvement) would also correlate more highly with the OEQ-O than the OEQ-E scales.

It was therefore decided that a sample of patients would be asked, in addition to completing the standard PROMs follow-up questionnaires (6 months post procedure for joint replacements, 3 months following VV surgery), also to complete the eleven items of the OEQ. This sample would be a consecutive series of patients recruited from all provider trusts over a specific time period to achieve the target. Patients were randomly selected from within a normal week's data feed to obtain coverage across providers and across all three procedures. Data were collected during October and November 2011.

The second study examined the performance of the OEQ alongside an established survey, the NHS inpatient survey, a regular survey of adults receiving inpatient care from NHS hospitals. The survey at the time of this study comprised 34 questions about different aspects of experiences of services as an inpatient. The questions had been grouped a priori by Department of Health analyses and Picker Europe¹⁰ into 12 overall domains with domain cores transformed to range 0–100.

The rationale of the second study was to see whether, when administered alongside the NHS inpatient survey of patient experiences, the OEQ-E would correlate more highly than the OEQ-O with other measures of patient experience. It was also hypothesized that the OEQ-E would correlate more strongly with relational measures in the inpatient survey compared to other aspects of care, such as cleanliness or access.

For this second study, three contrasting trusts agreed to participate: a teaching hospital in London, a general hospital from northern England, and a general hospital trust from the Midlands. The OEQ was presented as an additional, separate questionnaire within the envelope enclosing the inpatient survey. Data were collected during September 2011–January 2012.

Measures

A range of measures were included in relation to the OEQ. In the first study, the OHS has 12 items relating to functioning and pain with patients undergoing hip replacement with

a summed score reflecting the severity of problems related to their hip. Responses are obtained on a 5-point Likert scale. Scores range from 12 to 60 with high scores representing better health. The OKS has 12 items relating to functioning and pain with patients undergoing hip replacement with a summed score reflecting the severity of problems related to their hip. Responses are obtained on a 5-point Likert scale. Scores range from 12 to 60 with high scores representing better health. The AVVQ contains 13 items relating to problems of venous disease of the legs. The questionnaire is scored from 0 to 100, where 0 represents a patient with no evidence of VV and 100 represents the most severe problems associated with VV. The EQ-5D is a generic, preference based health status measure with items: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Scores are obtained on three levels: no problems, some problems, extreme problems. An EQ-5D health state may be converted to a single summary index.

Change scores for measures administered before and after surgery (OHS, OKS, AVVQ, EQ-5D) were calculated as simple differences between scores.

In the second study, the NHS inpatient experience survey comprises of 34 items assessing 12 domains, scores range from 0 to 10 with lower scores indicating scope for improvement.

Statistical analysis

Data were analyzed using SPSS version 20. Chi square was used for categorical variables and Fisher's one-way analysis of variance [ANOVA] for continuous variables. McNemar's test was used to examine the differences between the baseline and follow-up questionnaires for dichotomous variables. Pearson's correlation coefficients were used to calculate relationships between variables. Internal consistency, the extent to which items comprising a scale measure the same construct (eg, homogeneity of items in a scale); was assessed by Cronbach's alphas. Cronbach's alphas 0.70–0.90 are considered adequate.¹⁵

Results

Study one: OEQ and NHS national PROMs program

Response rates and respondent characteristics

In total, a sample of 720 patients included in the national PROMs program were sent a baseline questionnaire by post. Twenty-eight people were not eligible for inclusion in the study due to canceled operations, the patient having died or being unobtainable through moving address. Therefore, 692 people were eligible for inclusion in the study.

Overall, after surgery 490 people (71%) returned a follow-up PROMs questionnaire together with the OEQ. High response rates for the return of the PROMs and the OEQ were observed for the hip and knee groups (80% and 83%, respectively); however, the response rate for the VV group was much lower (50%). Only three people (0.4%) directly refused to participate in the study. Amongst those who returned the OEQ, the proportion of missing responses for all items in both sub-scales was very low (range from 1.4% to 0.4%), suggesting the items in the sub-scales appear acceptable to respondents.

Respondent characteristics, general health, and satisfaction with operation

Respondent characteristics are shown in Table 1. Respondents in the VV group were significantly younger than in the hip or knee groups ($F=52.36$, $P=0.001$). There were no differences between the groups in regards to sex, living arrangements, overall assessment of their operation or problems with their

condition in comparison to before the operation. Respondents in the VV group more frequently reported having their condition for 6 or more years (chi-square =90.09, $P=0.001$) and having received previous treatment for their condition (chi-square =61.65, $P=0.001$). In the baseline questionnaire, approximately half of all respondents reported having a disability prior to treatment. Significantly less people in the VV group reported having a disability (chi-square =80.16, $P=0.001$). There were no differences between the three groups in terms of general health at baseline or follow-up, however, all groups (hip, knee, VV) reported significant improvement in their general health at follow-up in comparison to baseline ($P<0.05$).

Internal consistency reliability

The internal consistency reliability (Cronbach's alpha) for the two sub-scales of the OEQ: OEQ-O 0.86 (hip and VV) and 0.92 (knee) with overall 0.89. For the OEQ-E 0.84 (VV) to 0.90 (hip) and overall 0.88. This suggests homogeneity of the scales.

Table 1 Respondent characteristics for the three groups and combined across groups

Variable	Hip group		Knee group		Vein group		Combined	
	No	%	No	%	No	%	No	%
Age								
"Mean" age	69 years		71 years		58 years		67 years	
SD (years)	10.95		8.87		13.91		12.16	
Range (years)	30–92		44–87		17–96		17–96	
Sex								
Male	72	39.3%	79	42.5%	52	46.4%	203	42.2%
Female	111	60.7%	107	56.9%	60	53.6%	278	56.7%
Q1 general health								
Excellent	8	4.6%	1	0.6%	8	7.4%	17	3.7%
V Good	58	33.5%	36	20.0%	42	37.5%	136	29.2%
Good	69	39.9%	98	54.4%	45	40.2%	212	45.6%
Fair	30	17.3%	36	20.0%	17	14.5%	83	17.9%
Poor	8	4.6%	9	5.0%	–	0%	17	3.7%
Q2 general health								
Excellent	15	8.5%	6	3.2%	12	10.3%	33	7.0%
V Good	75	42.6%	57	30.3%	44	40.0%	176	37.3%
Good	63	35.8%	82	43.6%	38	34.5%	183	38.9%
Fair	21	11.9%	34	18.1%	15	13.6%	70	14.9%
Poor	2	1.1%	6	3.2%	1	0.9%	9	1.9%
Overall assessment of operation								
Excellent	64	36.8%	33	17.8%	27	24.3%	124	26.4%
V Good	65	37.4%	71	38.4%	40	36.0%	176	37.4%
Good	32	18.4%	54	29.2%	29	26.1%	115	24.5%
Fair	9	5.7%	19	10.3%	11	9.9%	40	8.5%
Poor	3	1.7%	8	4.3%	4	3.6%	15	3.2%
Assessment of problems compared to before operation								
Excellent	151	85.8%	129	69.4%	74	66.7%	354	74.8%
V Good	15	8.5%	36	19.4%	20	18.0%	71	15.0%
Good	3	1.7%	8	4.3%	10	15.3%	21	4.4%
Fair	5	2.8%	8	4.3%	4	6.3%	17	3.6%
Poor	2	1.1%	5	2.7%	3	2.7%	10	2.1%

Notes: Q1 – Baseline measurement; Q2 – post surgery.

Abbreviations: V, very; SD, standard deviation.

Descriptive statistics for the OEQ-O and OEQ-E

The descriptive statistics for both the OEQ-O and OEQ-E are shown in Table 2. Mean scores for OEQ-O and OEQ-E for all groups were similar.

Construct validity (OEQ and disease-specific measures)

It was hypothesized that the OEQ-O would have stronger correlations with the disease-specific and EQ-5D change scores (both measures of outcomes) than the OEQ-E. It was further hypothesized that the OEQ-O would have strong correlations with global judgment items (eg, overall assessment of operation, perceived improvement of condition). The correlations between the OEQ-O and the change score for the disease-specific measures were strong (OHS 0.50; OKS 0.68, both $P < 0.001$), and higher than the corresponding OEQ-E change score correlations (OHS 0.23; OKS 0.29, both $P < 0.05$). The association between the OEQ-O and the OEQ-E was not significant for the AVVQ change score ($-0.18, 0.04$).

Construct validity (OEQ and EQ-5D)

The pattern of correlations between the OEQ and the EQ-5D was similar to those observed for the disease-specific measures. The OEQ-O was moderately correlated with the EQ-5D change scores for each condition (hip 0.38; knee 0.30; VV 0.34, all $P < 0.01$), and demonstrated higher associations compared to the equivalent correlations of OEQ-E (hip 0.16, $P < 0.05$; knee 0.14; VV -0.01 , both not significant).

OEQ and PROMs single item outcomes

The degree to which the OEQ scales performed in regards to similar constructs was assessed by comparing the

OEQ scales with the single-item “overall assessment of operation” and “perceived improvement after operation” questions included in the PROMs questionnaire. The OEQ-O was more strongly associated with the “overall assessment of operation” in comparison to the OEQ-E (Table 3).

Similarly the OEQ-O was more strongly associated with the “improvement after operation” item in comparison to the OEQ-E (Table 4).

Study 2: OEQ and NHS inpatient survey

Data were available from a total of 586 patients from the three sites who returned the OEQ and inpatient survey. The wide range in age (17–98 years) and length of stay (1–112 days) are reflected in large standard deviations for these variables (standard deviation 17.6 and 9.8 respectively). A total response rate of 32.6% was achieved (Table 5). Levels of missing data for all items from returned questionnaires were very low, (range from 1% to 4.1% per item).

Internal consistency reliability

Cronbach's alpha values were calculated at 0.86 for the OEQ-O domain and 0.89 for the OEQ-E domain providing further support of homogeneity of the scales.

Construct validity

Table 6 presents correlations between the two domains of the OEQ and the 12 a priori determined core domains from the NHS inpatient survey (34 items). As hypothesized, all correlations of NHS inpatient survey domains with the OEQ-E domain are greater than those with the OEQ-O domain, thus providing evidence for construct validity of the OEQ. Lowest correlations were observed for OEQ-E with NHS inpatient survey items addressing cleanliness and access and waiting, as hypothesized.

Discussion

The data presented here provide preliminary evidence of the potential of the OEQ to assess patients' views about how they experience care and the outcomes of that care.

Table 3 OEQ-O and OEQ-E correlations with “overall assessment of operation”

	Hip group	Knee group	Vein group	Combined
OEQ-O	0.77	0.71	0.77	0.75
OEQ-E	0.55	0.34	0.35	0.40

Note: All associations significant at $P = 0.001$.

Abbreviations: OEQ-O, Outcomes and Experiences Questionnaire-Outcomes; OEQ-E, Outcomes and Experiences Questionnaire-Experiences.

Table 2 OEQ-O and OEQ-E scale mean, standard deviation, median, and range

Variable	Hip group	Knee group	Vein group	Combined
OEQ-O				
Mean	17.19	16.19	16.05	16.54
SD	3.24	4.36	3.99	3.91
Median	18	17	17	18
Range	0–20	0–20	4–20	0–20
OEQ-E				
Mean	15.24	15.23	15.83	15.37
SD	3.19	3.09	2.65	3.04
Median	16	16	17	16
Range	4–18	2–18	6–18	2–18

Abbreviations: SD, standard deviation; OEQ-O, Outcomes and Experiences Questionnaire-Outcomes; OEQ-E, Outcomes and Experiences Questionnaire-Experiences.

Table 4 OEQ-O and OEQ-E correlations with “improvement after operation”

	Hip group	Knee group	Vein group	Combined
OEQ-O	0.77	0.85	0.75	0.80
OEQ-E	0.41	0.30	0.11	0.27

Notes: All associations significant at $P=0.001$ unless otherwise stated; figures in italics are not significant associations.

Abbreviations: OEQ-O, Outcomes and Experiences Questionnaire-Outcomes; OEQ-E, Outcomes and Experiences Questionnaire-Experiences.

The study design did not provide a straightforward assessment of the acceptability of the OEQ because in both validation surveys the OEQ had to be sent with a pack containing other standard instruments so it is difficult to disaggregate effects of the OEQ. Response rates were typical of those prevailing for the existing surveys into which OEQ was inserted. However very few respondents in either survey completed the main standard survey instrument without also completing the OEQ. The rate of missing responses at the item level for OEQ was very low in both surveys suggesting that no items posed obvious difficulties for respondents. A possible weakness of the first validation study was the pragmatic need to administer the OEQ at the time required by the national PROMs program, (3 or 6 months after surgery). These may not have been optimal times to obtain reports of experiences of care, the salience of which may have diminished.

In both studies, the internal consistency of the two sub-scales was high, providing encouraging evidence of this aspect of reliability. Reproducibility could not be assessed in this study and needs to be assessed in any further evaluations.

The main focus of the two validation studies was to examine whether the OEQ produced responses consistent with respondents having distinct and distinguishable views in relation to experiences and outcomes. Absence of such evidence would suggest that respondents made more global or conflated judgments and would weaken the case for the OEQ.

In the context of the national PROMs program, the evidence was strikingly positive in relation to the two orthopedic surgical procedures. The OEQ-O sub-scale correlated very much more strongly than did the OEQ-E sub-scale with condition-specific change scores. For hip and knee replacement patients, judgments of the benefits of treatment captured in OEQ-O agreed strongly with changes over time in

Table 5 Response rates by site and total sample

	Returned (n)	Response rate
Site 1	225	21.0%
Site 2	235	37.5%
Site 3	126	39.2%
Total	586	32.6%

Table 6 Correlation coefficients between the OEQ domains and 12 a priori determined core domains from the NHS inpatient survey

NHS inpatient survey domains	OEQ-O	OEQ-E
Consistency and coordination of care	0.58**	0.68**
Treatment with dignity and respect	0.43**	0.62**
Involvement	0.51**	0.63**
Doctors	0.54**	0.63**
Nurses	0.40**	0.57**
Cleanliness	0.30**	0.37**
Pain control	0.40**	0.51**
Access and waiting	0.23**	0.26**
Safe, high quality, coordinated care	0.37**	0.47**
Better information, more choice	0.40**	0.45**
Building better relationships	0.44**	0.58**
Clean, comfortable, friendly place to be	0.42**	0.58**

Note: **Correlation significant at 0.01.

Abbreviations: OEQ-O, Outcomes and Experiences Questionnaire-Outcomes; OEQ-E, Outcomes and Experiences Questionnaire-Experiences; NHS, National Health Service.

those aspects of health status that the surgical procedures are intended to improve. On the other hand patients' views about how they were cared for, as captured through OEQ-E, were only modestly related to changes in health targeted by the two orthopedic surgical procedures.

This is very encouraging evidence of the absence of halo effects or global patterns of judgment, and is consistent with the evidence of Black et al who also report positive but modest correlations between patients' reports of outcomes and experiences in the national PROMs program.¹⁶

The patterns of correlations of OEQ and change scores for EQ-5D were as hypothesized. There were significant correlations between OEQ-O and EQ-5D change scores compared with weaker or non-significant correlations for OEQ-E and EQ-5D change scores for all three surgical groups. The lower correlations of OEQ-O with EQ-5D compared with condition-specific measures were predicted because EQ-5D partly taps aspects of general health not expected to be directly impacted by surgery.

OEQ was examined in relation to the two single retrospective items in the PROMs program, overall assessment and perceived improvement. A clear pattern, very consistent with hypotheses, emerged of OEQ-O correlating much more strongly than OEQ-E with both items for each of the three surgical groups and for the overall combined sample.

The patterns of correlations observed for the AVVQ in patients receiving VV surgery were less supportive of the hypotheses explored in this paper. From other analyses beyond the scope of this paper, it was inferred that the AVVQ was a far less responsive measure of outcome for VV surgery and less helpful as a benchmark for evaluating OEQ.

The examination of validity in the context of the NHS inpatient survey was in many respects the mirror image of the first study. It was hypothesized that the OEQ-E sub-scale would correlate more strongly than the OEQ-O sub-scale with domains of patient experience measured in the existing NHS survey instrument. This was indeed consistently observed. Moreover it was also hypothesized that the OEQ-E would correlate less strongly with aspects of experience that were non-relational. This was also observed with OEQ-E correlating least strongly with domains of patient experience such as cleanliness and access that are outside of core clinical domains of patient experience.

Thus there was supportive evidence for validity in the second validation study. However the pattern of differences in correlations of OEQ-O and OEQ-E with the existing measures was not as striking as in the PROMs elective surgery study. Indeed the OEQ-O sub-scale did consistently correlate significantly (although less strongly) with domains of patient experience. It is difficult to interpret this result with the information available from the inpatient survey onto which this validation study was added. Limited data are available regarding reasons for admission and, for example, some patients may have been admitted for a very wide range of procedures, sometimes for assessment or investigative procedures in addition to or instead of treatment. Overall the possibility that patients were making global judgments influenced by halo-effects or judgments conflating distinct domains of experience and outcome is less easy to exclude with the results observed in the second validation study.

Overall the evidence of the performance of OEQ in two contrasting acute hospital contexts is encouraging, with broadly positive evidence of its acceptability and capacity to elicit distinct judgments of some key aspects of outcomes and experiences of health care. The instrument does require further testing in different contexts and populations; specifically in a wider range of interventions. Currently the OEQ is being tested in the context of an evaluation of care planning for individuals with long-term conditions in the community, separate studies of musculoskeletal services, and a comparative study of hospital providers.

If evidence accumulates that the OEQ is an acceptable, reliable and valid instrument, it offers some important advantages to capturing patients' views. It is short and likely to be feasible and practical to collect, process, and report. It addresses views about a substantial range of domains of health care that it is reasonable to think matter to patients. It can be used in contexts where well-validated PROMs do not yet exist or where they are not, in their traditional form

of administration pre- and post-intervention, feasible. Above all by bringing together into one arena the two hitherto largely unrelated domains of outcomes and experiences, the OEQ will contribute to the process of putting onto an equal footing these two domains. Furthermore, this supports a person-centered approach to measurement as a potential to improve quality of care.¹⁷

There are some potential disadvantages of the approach to measurement entailed by OEQ. The main disadvantage is that the OEQ depends on retrospective items with no specific content about symptoms or other aspects of health. Standard PROMs, which typically do contain patients' reports of problems familiar and relevant to the clinician, are just beginning to be accepted as part of the landscape of outcome measurement. As has recently been argued by Grosse Frie et al,¹⁸ even if shown to have substantial validity, items largely focused around retrospective judgment of benefit may be challenging for key health professional audiences because of their distance from instruments focused on health-related content and change scores between pre- and post-intervention.

Conclusion

Overall evidence is provided to support the view that it is feasible to bring together into a single instrument questions addressing key concerns of patients regarding their experiences of how they were cared for and how they judge the outcomes of that care. Further testing in a wider range of contexts is needed to support the uptake of the OEQ as a summary measure of patients' key views.

Acknowledgments

Funding for this study was provided by the Department of Health Policy Research Programme and National Institute of Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC -Oxford).

We would like to acknowledge and thank our patient partners who provided valuable comments during development of items for the OEQ.

In the first study, consent to participate was obtained from patients as described in the national guidance documentation for the PROMs programme: http://www.hscic.gov.uk/media/1537/A-Guide-to-PROMs-Methodology/pdf/PROMs_Guide_V8.pdf; <http://www.hscic.gov.uk/catalogue/PUB14574/final-proms-eng-apr12-mar13-fin-report-v1.pdf>

In the second survey, consent was implied by patients' willingness to complete and return a questionnaire, having received an explanation for the survey. National ethics

approval for the method of consenting had been obtained by the survey company, Picker Institute Europe: <http://www.nhssurveys.org/survey/1097>.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Coulter, A, Jenkinson C. European patients' views on the responsiveness of health systems and healthcare providers. *Eur J Public Health*. 2005;15(4):355–360.
2. Schoen C, Osborn R, Doty MM, Bishop M, Peugh J, Murukutla N. Toward higher-performance health systems: adults' health care experiences in seven countries. *Health Aff (Millwood)*. 2007;26(6):w717–w734.
3. Department of Health. *Equity and Excellence: Liberating the NHS*. London: Department of Health; 2010. The Stationery Office Limited. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213823/dh_117794.pdf. Accessed June 7, 2015.
4. Department of Health. *NHS Outcomes Framework 2012/13*. London: Department of Health; 2011. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213711/dh_131723.pdf. Accessed June 7, 2015.
5. Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH. Comparative measurement sensitivity of short and longer health status instruments. *Med Care*. 1992;30(10):917–925.
6. Lloyd H, Jenkinson C, Hadi M, Gibbons E, Fitzpatrick R. Patient reports on the outcomes of treatment: a structured review. *Health Qual Life Outcomes*. 2014;12:5.
7. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, D.C. National Academy Press; 2001. Available from: http://www.nap.edu/html/quality_chasm/reportbrief.pdf. Accessed June 7, 2015.
8. Coulter A, Fitzpatrick R, Cornwell J. *The Point of Care Measures of Patients' Experience in Hospital: Purpose, Methods and Uses*. Kings fund, London; 2009. Available from: <http://www.kingsfund.org.uk/sites/files/kf/Point-of-Care-Measures-of-patients-experience-in-hospital-Kings-Fund-July-2009.pdf>. Accessed June 7, 2015.
9. Picker Institute [homepage on the Internet]. Picker Principles of Patient-Centered Care;1987. Available from: <http://pickerinstitute.org/about/picker-principles/>. Accessed September 2, 2014.
10. Murrells T, Robert G, Adams M, Morrow E, Maben J. Measuring relational aspects of hospital care in England with the 'Patient Evaluation of Emotional Care during Hospitalisation' (PEECH) survey questionnaire. *BMJ Open*. 2013;3(1).pii:e002211.
11. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg Br*. 1996;78(2):185–190.
12. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br*. 1998 Jan;80(1):63–69.
13. Garratt AM, Macdonald LM, Ruta DA, Russell IT, Buckingham JK, Krukowski ZH. Towards measurement of outcome for patients with varicose veins. *Qual Health Care*. 1993;2(1):5–10.
14. Szende A, Oppe M, Devlin N. *EQ-5D Value Sets: Inventory, Comparative Review and User Guide*. Dordrecht, The Netherlands: Springer; 2007.
15. Nunnally JC, Bernstein IH. *Psychometric Theory*. McGraw-Hill Series in Psychology. 3rd ed. McGraw-Hill Higher Education; 1993.
16. Black N, Varaganum M, Hutchings A. Relationship between patient reported experience (PREMs) and patient reported outcomes (PROMs) in elective surgery. *BMJ Qual Saf*. 2014;23(7):534–542.
17. The Health Foundation. *Helping Measure Person-Centred Care*. London: The Health Foundation; 2014. Available from: <http://www.health.org.uk/public/cms/75/76/313/4697/Helping%20measure%20person-centred%20care.pdf?realName=Lnl7Fn.pdf>. Accessed June 7, 2015.
18. Grosse Frie K, van der Meulen J, Black N. Single item on patients' satisfaction with condition provided additional insight into impact of surgery. *J Clin Epidemiol*. 2012;65(6):619–926.

Supplementary material

The Outcomes and Experiences Questionnaire

Q1. How helpful has your most recent visit to hospital been in dealing with the problem(s) you came to hospital for?

- ☐₁ Extremely helpful
- ☐₂ Very helpful
- ☐₃ Helpful
- ☐₄ A little helpful
- ☐₅ Not at all helpful
- ☐₆ Problem(s) completely cured

Q2. How would you now rate the problem(s) you recently came to hospital for?

- ☐₁ Much better
- ☐₂ A little better
- ☐₃ The same
- ☐₄ A little worse
- ☐₅ Much worse

Q3. How helpful was your most recent visit to hospital in helping you manage any aspects of the problem(s) that continued after you left hospital?

- ☐₁ Extremely helpful
- ☐₂ Very helpful
- ☐₃ Helpful
- ☐₄ A little helpful
- ☐₅ Not at all helpful
- ☐₆ No problems remained; problem(s) completely cured

Q4. How would you rate your health now as a result of your hospital visit?

- ☐₁ Much better
- ☐₂ A little better
- ☐₃ The same
- ☐₄ A little worse
- ☐₅ Much worse

Q5. When you had important questions to ask staff, did you get answers that you could understand?

- ☐₁ Yes, always
- ☐₂ Yes, most of the time
- ☐₃ Yes, some of the time
- ☐₄ No, never

Q6. How helpful was the information you were given about your treatment and condition at your most recent hospital visit?

- ☐₁ Extremely helpful
- ☐₂ Very helpful
- ☐₃ Helpful
- ☐₄ A little helpful
- ☐₅ Not at all helpful
- ☐₆ I was not given information but would have liked some
- ☐₇ I did not need any information

Q7. Were you involved as much as you wanted to be in decisions about your care and treatment at your most recent hospital visit?

- ☐₁ As much as I wanted to be
- ☐₂ Less than I wanted to be
- ☐₃ Not at all although I wanted to be
- ☐₄ Not at all and I did not wish to be
- ☐₅ I was more involved than I wanted to be

Q8. How much did hospital staff respond to your individual needs during your most recent hospital visit?

- ☐₁ At all times
- ☐₂ Most of the time
- ☐₃ Some of the time
- ☐₄ None of the time

Q9. Were you able to discuss any worries and fears with staff during your most recent hospital visit?

- ☐₁ As much as I wanted
- ☐₂ Most of the time
- ☐₃ Some of the time
- ☐₄ Not at all, but would have liked to
- ☐₅ I did not have any worries or fears

Q10. Did the different people treating and caring for you work well together to give you the best possible care?

- ☐₁ Yes, always
- ☐₂ Yes, most of the time
- ☐₃ Yes, some of the time
- ☐₄ No never
- ☐₅ Don't know

Q11. Overall, how would you rate the outcome of your most recent visit to hospital?

- ☐₁ Excellent
- ☐₂ Very good
- ☐₃ Good
- ☐₄ Fair
- ☐₅ Poor

Patient Related Outcome Measures

Dovepress

Publish your work in this journal

Patient Related Outcome Measures is an international, peer-reviewed, open access journal focusing on treatment outcomes specifically relevant to patients. All aspects of patient care are addressed within the journal and practitioners from all disciplines are invited to submit their work as well as healthcare researchers and patient support groups.

Submit your manuscript here: <http://www.dovepress.com/patient-related-outcome-measures-journal>

The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.