BJA

British Journal of Anaesthesia, 116 (2): 241–8 (2016)

doi: 10.1093/bja/aev413
Clinical Practice

Measuring quality of recovery-15 after day case surgery

M. Chazapis[1,2,3,4,*], E. M. K. Walker[1,2,3,4], M. A. Rooms[2,3], D. Kamming[2,3] and S. R. Moonesinghe[1,2,3,4,5]

[1]UCLH Surgical Outcomes Research Centre, Department of Applied Health Research, University College London, UK, [2]Centre for Anaesthesia, University College London, UK, [3]Department of Anaesthesia and Perioperative Medicine, University College Hospital, London, UK, [4]National Institute for Academic Anaesthesia's Health Services Research Centre, Royal College of Anaesthetists, London, UK, and [5]UCL Hospitals NIHR Biomedical Research Centre, University College London Hospitals NHS Foundation Trust, London UK

*Corresponding author. E-mail: m.chazapis@gmail.com

Abstract

Background: 'Quality of recovery' scores are patient-reported outcome measures evaluating recovery after surgery and anaesthesia. However, they are not widely used in the clinical or research setting. The Quality of Recovery-15 (QoR-15) is a recently developed, psychometrically tested and validated questionnaire.

Methods: We conducted a prospective study of all adult patients undergoing orthopaedic day case surgery over a period of six months (June 2013–November 2013). Patients completed the QoR-15 score preoperatively, and then were asked to repeat the score by telephone at 24 h, 48 h and seven days after surgery.

Results: 633 patients from a possible 714 (89%) completed the preoperative questionnaire and data from 437 patients who completed scores at all four time points were analysed. Most patients returned to their preoperative score by 48 h, and had exceeded it by seven days. Construct validity was supported by a negative correlation with duration of surgery and total inpatient opioid use. There was also excellent internal consistency (Cronbach's alpha 0.80–0.83).

Conclusions: The QoR-15 is a clinically acceptable and feasible patient-centred outcome measure after day case surgery. The score demonstrated good validity, reliability and responsiveness. However, measurement of the QoR-15 score on the day of surgery may not provide a true baseline value. We suggest one follow-up call at 48 h would enable an adequate patient-centred assessment of postoperative recovery after day case orthopaedic surgery.

Key words: ambulatory surgical procedures E04.030; anesthesia E03.155; health care N05.700, quality assessment

Day case surgery is an expanding speciality. An increasing number of patients are being considered suitable for more complex surgery.[1] The challenge is to maintain the quality of care and improve patient outcomes within this type of healthcare delivery.[2]

Assessing postoperative patient recovery has traditionally focused on outcome measures of morbidity, mortality, physiological changes and re-hospitalization rates.[3] These are important and should be measured, but these data represent only one aspect of a patient's recovery. A patient's ability to resume normal activities after surgery and anaesthesia is an important indicator of a successful perioperative experience.[4] Measuring the quality of recovery (QoR) from a patient's perspective requires an assessment of multiple patient-centred outcomes.[4]

Multiple quality of recovery tools have been developed.[5 6] However, existing studies have focused predominantly on inpatient surgery rather than a day case setting. The QoR-15 is a recently developed and validated short-form postoperative QoR score (Supplementary material).[7] Fifteen questions assess five

Accepted: October 25, 2015

Editor's key points

- The QoR score is an established tool to measure quality of recovery after surgery, but there are no data on its use for day case surgery.
- This single-centre study evaluated the QoR score in adults undergoing day case orthopaedic surgery.
- The QoR score was found to be feasible, reliable and consistent.
- However, scores taken within 24 h before surgery may not represent a 'true' baseline because of fatigue, anxiety or other factors.

domains of patient-reported health status: pain, physical comfort, physical independence, psychological support and emotional state. The 11-point numerical rating scale leads to a minimum score of 0 (very poor recovery) and a maximum score of 150 (excellent recovery).

The aim of this study is to test the acceptability and feasibility of using the QoR-15 score as a patient-reported outcomes measure after day case surgery and anaesthesia, and to identify when we should be administering the QoR-15 score.

Methods

After approval by the local research ethics committee and designation as a service evaluation, we conducted a prospective, observational, cohort study of all adult patients undergoing day case orthopaedic surgery over a period of six months (June 2013–December 2013) in University College Hospital, London. Patients were sequentially recruited and written consent obtained. Day case surgery was defined as surgical procedures not requiring a planned overnight hospital stay.

Patients were excluded if they had a known history of alcohol or drug abuse, a psychiatric disturbance precluding complete cooperation, poor English understanding, or they were aged less than 18 yr old.

Eligible patients were approached and given a patient information leaflet on the day of surgery. Patients filled in the preoperative QoR-15 questionnaire (Supplementary material) before their surgery. This completed questionnaire was a baseline measure of health status over the previous 24 h. The patients were then telephoned by research nurses, and repeated the same questionnaire at 24 h, 48 h and seven days after surgery.

Additionally, anaesthetic and recovery staff collected patient characteristics and perioperative data. This included: age; gender; ethnicity; ASA physical status; BMI; extent, type, and duration of surgery; type of anaesthesia and analgesia use.

The QoR-15 was psychometrically evaluated using data collected from patients who responded at all four time intervals.[8 9] This included:

(i) Acceptability and Feasibility - These were assessed by the:
   1. Patient recruitment rate
   2. Successful completion rate of the questionnaire at all four time points
   3. Time taken to complete the questionnaire in a subset of patients (n=50)
   4. Which patients were more likely to respond or not respond to the questionnaire at any of the postoperative time points

(ii) Validity – This describes accuracy of the questionnaire. Construct validity was explored, investigating associations between the QoR-15 at 24 h with age, gender, duration of surgery and total opioid use (by calculating total morphine equivalents).

(iii) Reliability – This describes consistency of questionnaire responses and was assessed using:
   1. Internal consistency: This assesses the consistency of results across items within a test.
   2. Inter-item correlation matrix: This assesses the correlation of individual items within a test.
   3. Inter-dimension and Item-to-total dimension correlations: These assess the correlation and consistency of the individual dimensions within a test.

(iv) Responsiveness – This describes the questionnaire's ability to detect change at a group and individual level and was assessed using:
   1. Cohen effect size, calculated as the average change scores (from pre-test to post-test) divided by the sd at baseline.[10]
   2. Standardised response mean, calculated as the change of scores divided by the sd of the change scores.[10]

Statistical analysis

The sample size of this study was guided by previous studies, as power calculations cannot be reliably determined with correlation analysis.[7] Data are presented as mean (SD), median (interquartile range), number (%) or 95% CI. All percentages are rounded up to the nearest integer.

Continuous data were tested for normality using the Shapiro-Francia and Shapiro Wilk normality tests; No variables were normally distributed.

To compare the patients who completed all three of the postoperative questionnaires against the patients that did not, the distributions of gender, ASA, BMI, smoking, presence of comorbidities, type of surgery (upper limb or lower limb), use of regional anaesthesia, type of general anaesthesia, and analgesic use were compared using χ^2 tests. The distributions of age, duration of surgery, and total opioid use, were compared using two sample Kolmogorov-Smirnov tests.

Associations were measured using Spearman rank correlation coefficient (ρ) presented to two decimal places. Statistical significance was set at a P value of $\leq$0.05 and where necessary, Bonferroni's correction was used to adjust for multiple comparisons with a corrected P value (P') of <0.05. The non-parametric Wilcoxon rank sum test (Mann–Whitney U-test) (z) and Kruskal Wallis tests were also used to compare QoR-15 scores.

Internal consistency was measured using Cronbach's alpha.[11] Inter-dimension and item-to-total dimension correlation coefficients and average inter-item covariances were also measured. Comparisons between the total QoR-15 scores at different time points were made using Friedman's non-parametric anova, followed by Wilcoxon's matched pairs test in case of significance. Bonferroni's correction was used to adjust for multiple comparisons with a corrected P value (P') of <0.05.

All statistical analyses were performed using STATA/IC for Mac v12.1 (StataCorp LP, Texas, USA).

Results

Over the study period of six months, a total of 714 patients were eligible for inclusion. The study flowchart is presented in Fig. 1. 633 evaluable patients completed the preoperative and at least one postoperative questionnaire giving a recruitment rate of 89%. Clinical characteristics of included and excluded patients

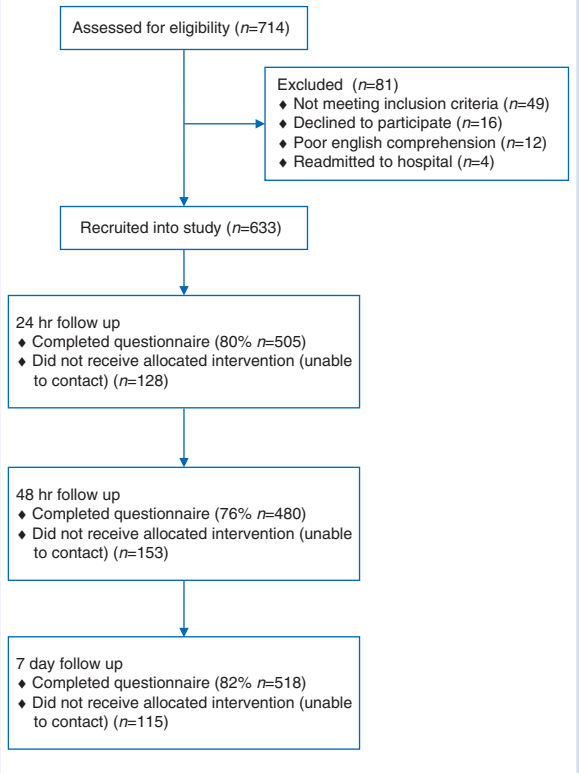

Fig 1 Study Flowchart.

Table 1 Patients' characteristics. Number (%) or median (interquartile range) unless otherwise stated

	Included patients (n=633)	Patients who responded at all 4 time points (n=437)	Excluded patients (n=81)
Age, yr Range	18–88 (47)	18–85 (47)	18–74 (44)
Sex (M/F) (%M)	308/325 (49)	214/223 (49)	38/43 (47)
BMI (kg m^{-2}) n (%)			
<20	17 (3)	12 (3)	2 (3)
20–25	208 (33)	169 (39)	25 (31)
26–30	198 (31)	144 (33)	27 (33)
>30	99 (16)	73 (17)	12 (15)
>40	9 (1)	8 (2)	1 (1)
>50	8 (1)	6 (1)	0 (0)
not recorded	94 (15)	25 (6)	14 (17)
Ethnicity			
British	365 (58)	269 (62)	40 (49)
Any other White background	91 (13)	54 (12)	15 (19)
Black	24 (4)	17 (4)	4 (5)
Any other Black background	4 (1)	2 (0.5)	0 (0)
Asian	47 (7)	27 (6)	8 (10)
Mixed	16 (3)	11 (2.5)	1 (1)
Any other ethnic background	42 (7)	27 (6)	4 (5)
not recorded	44 (7)	30 (7)	9 (11)
Comorbidities			
Hypertension	73 (12)	57 (13)	11 (14)
Atrial Fibrillation	8 (1)	7 (2)	1 (1)
Angina	8 (1)	6 (1)	0 (0)
MI	7 (1)	7 (2)	2 (2)
COPD	12 (2)	7 (2)	1 (1)
Asthma	60 (9)	47 (11)	9 (11)
Renal disease	3 (0.5)	3 (0.7)	0 (0)
Neuropathy	13 (2)	9 (2)	2 (2)
Diabetes	32 (5)	23 (5)	7 (9)
Malignancy	1 (0.2)	1 (0.2)	0 (0)
Smoker	97 (15)	76 (17)	12 (15)
ASA			
I	365 (58)	257 (59)	49 (60)
II	214 (34)	144 (33)	24 (30)
III	29 (5)	23 (5)	3 (4)
IV	0 (0)	0 (0)	0 (0)
not recorded	25 (4)	13 (3)	5 (6)
Type of Orthopaedic Surgery			
Shoulder/Clavicle	130 (23)	100 (23)	16 (20)
Upper Limb/Other	36 (7)	23 (5)	4 (5)
Wrist/Hand	98 (16)	68 (16)	15 (19)
Upper Limb Total	264 (42)	191 (44)	35 (43)
Lower Limb/Other	50 (7)	33 (8)	4 (5)
Knee	132 (21)	88 (20)	19 (23)
ACL repair	24 (4)	13 (3)	4 (5)
Foot/Ankle	163 (26)	107 (2)	19 (23)
Lower Limb Total	369 (58)	241 (55)	46 (57)
Duration of surgery, min	55 (5–240)	55 (5–240)	60 (5–160)
Length of recovery stay, min	50 (5–320)	45 (5–260)	55 (5–310)

are presented in Table 1. There were no significant differences between the study patients and the excluded patients.

437 (69%) patients completed all three postoperative QoR-15 questionnaires. The mean time taken to complete the postoperative 24 h QoR-15 score in a subset of patients (*n*=50) was 2.6 (1–7) min. There were no differences in gender, ASA, age, BMI, smoking, or any of the recorded comorbidities other than hypertension (*P*=0.036) in those who completed all three of the postoperative questionnaires and those who did not. The study staff informally reported that patients were very happy to be contacted, found the questions easy to understand and overall felt it was a positive experience.

Box plots of total QoR-15 scores at each study time point are presented in Fig. 2. The percentage of patients achieving the highest possible QoR-15 score at the different time points were: ppreoperative (8%, *n*=34), 24 h (6%, *n*=25), 48 h (8%, *n*=36), and at seven days (17%, *n*=75). The data are negatively skewed; levels of skew are: −1.29 (24 h), −1.42 (48 h) and −2.4 (seven days).

There was no significant relationship between ethnicity and total QoR-15 scores at any of the four time points (using Spearman correlation: pre-op QoR-15 score, rho=-0.0274, *P*′=0.5699; QoR-15 score at 24 h, rho=0.0186, *P*′=0.7002; QoR-15 score at 48 h, rho=−0.0381, *P*′=0.4295; QoR-15 score at seven days, rho=−0.0641, *P*′=0.1836.

For each patient their total QoR-15 scores at each of the three postoperative time points were compared with their preoperative scores using the Wilcoxon signed-rank test. There was a difference between the preoperative total QoR-15 scores and those measured at 24 h (z=2.154, *P*=0.03) and seven days (z=−9.610, *P*=<0.001). Indeed, the seven day scores exceeded the preoperative 'baseline' scores. There was no difference between the preoperative total QoR-15 scores and those measured at 48 h (z=−1.197, *P*=0.23).

Construct validity was tested by comparing the total QoR-15 scores and patient's gender, age, duration of surgery and total

opioid use. There was no difference between pre-op total QoR-15 scores in men and women [132(17) *vs* 129(18), respectively, P=0.06] or at 48 h [133(16) *vs* 131(17), P=0.27] and seven days [140 (12) *vs* 137(15), P=0.07]. Men had higher total QoR-15 scores at 24 h [130(18) *vs* 126(19), P=0.02]. There was a negative correlation between age and the preoperative total QoR-15 score, which was statistically significant, ρ=−0.17, P'=<0.001; but this was not present at any of the postoperative time points.

There were negative correlations between the total QoR-15 score and length of surgery, at 24 h (ρ=−0.13, P'=0.006) at 48 h (ρ=−0.13, P'=0.009), and 7 days (ρ=−0.13, P'=0.007). There was a negative relationship between the total QoR-15 score and total inpatient opioid use at 24 h (ρ=−0.16, P'=0.002), 48 h (ρ=−0.12, P'=0.02), but not at seven days (ρ=−0.09, P'=0.06).

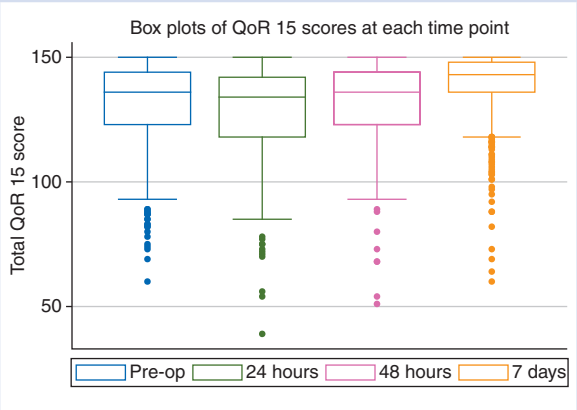

Fig 2 Box plots demonstrating median and IQR of total QoR-15 scores (437 patients) at each perioperative time point: preoperative, 24 h, 48 h and seven days after surgery. Whiskers demonstrate 5th and 95th percentiles.

Reliability indices measured using Cronbach's alpha (α) for the 437 patients for whom data were available at all time points, were high (>0.80) for all time points: α=0.83 for total preoperative QoR-15 scores, α=0.81 at 24 h, α=0.80 at 48 h and α=0.83 at seven days. The inter-item correlation matrix at 24 h is shown in Table 2. Inter-dimension and item-to-total dimension correlation coefficients at 24 h are included in Table 3.

Responsiveness was calculated using the Cohen effect size and standardized response means (SRM), included in Table 4 for the three postoperative time points. Cohen effect sizes of 0.2, 0.5 and 0.8 correspond to small, medium and large changes in quality of recovery scores. In this population, the total QoR-15 score had a Cohen effect size of 0.31 at 24 h, 0.30 at 48 h and 0.37 at seven days. Standardized response means of the total QoR-15 score were 0.21 at 24 h, 0.22 at 48 h and 0.33 at seven days.

Discussion

We have found the QoR-15 to be a clinically acceptable and feasible outcome measure after day case surgery. It demonstrates good validity, reliability and responsiveness. We suggest one follow-up call at 48 h would enable an efficient and clinically useful patient-centred assessment of postoperative recovery.

This is the largest study using the QoR-15 questionnaire to date; furthermore, only 21 patients evaluated in the original QoR-15 study underwent day case surgery. Few psychometric evaluations of QoR scores in the day surgery setting exist. Idvall and colleagues[12] psychometrically evaluated 399 patients, who fully completed a modified 29-item QoR-40 after day case surgery. Bost and colleagues[13] assessed the 8-item Short-Form Health Survey (SF-8) and the QoR-40 in 154 patients after day case anterior cruciate ligament repair surgery.

The high recruitment and response rate indicate that the QoR-15 is an acceptable and feasible outcome measure for day case patients. Acceptability of patient-based outcome measures is important to ensure high return rates and results in less bias from

Table 2 Inter-item Correlation Matrix for the QoR-15 at 24 h postoperatively (437 patients). Quality of recovery (QoR)-15 items: 1=able to breathe easily; 2=been able to enjoy food; 3=feeling rested; 4=have had a good sleep; 5=able to look after personal toilet and hygiene unaided; 6=able to communicate with family or friends; 7=getting support from hospital doctors and nurses; 8=able to return to work or usual home activities; 9=feeling comfortable and in control; 10=having a feeling of general well-being; 11=moderate pain; 12=severe pain; 13=nausea or vomiting; 14=feeling worried or anxious; 15=feeling sad or depressed. Inter-item Correlation Matrix for the QoR-15 at 24 h postoperatively (437 patients)

QoR-15 question number	Total QoR-15 score	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Breathing	0.33	—													
2. Food	0.59	0.18	—												
3. Rest	0.64	0.17	0.43	—											
4. Sleep	0.64	0.11	0.38	0.63	—										
5. Hygiene	0.48	0.17	0.24	0.23	0.23	—									
6. Communication	0.34	0.24	0.19	0.13	0.18	0.22	—								
7. Support	0.18	0.02	0.08	0.003	0.14	−0.02	0.12	—							
8. Return to work	0.62	0.10	0.26	0.27	0.31	0.37	0.15	0.02	—						
9. Feeling in control	0.71	0.18	0.38	0.44	0.40	0.23	0.19	0.05	0.42	—					
10. Well-being	0.70	0.24	0.42	0.49	0.40	0.20	0.21	0.07	0.34	0.72	—				
11. Moderate pain	0.50	0.07	0.09	0.21	0.20	0.16	0.09	0.05	0.32	0.35	0.23	—			
12. Severe pain	0.64	0.21	0.33	0.35	0.38	0.25	0.19	0.11	0.30	0.38	0.31	0.18	—		
13. Nausea/vomiting	0.41	0.17	0.41	0.14	0.17	0.09	0.22	0.13	0.15	0.14	0.12	0.13	0.26	—	
14. Anxiety	0.60	0.26	0.24	0.26	0.21	0.17	0.20	0.09	0.23	0.38	0.47	0.22	0.38	0.21	—
15. Depressed	0.52	0.18	0.16	0.24	0.19	0.22	0.19	0.05	0.18	0.31	0.45	0.15	0.30	0.10	0.67

Table 3 Inter-dimension and item-to-total dimension correlation coefficients. Inter-dimension and item-to-total dimension correlation coefficients calculated at 24 h, 48 h and seven days after surgery (437 patients)

QoR-15 Question Number	24 h Inter-dimension correlation coefficient	24 h Item-to-total dimension correlation coefficient	24 h Inter-item Cronbach alpha	48 h Inter-dimension correlation coefficient	48 h Item-to-total dimension correlation coefficient	48 h Inter-item Cronbach alpha	seven days Inter-dimension correlation coefficient	seven days Item-to-total dimension correlation coefficient	seven days Inter-item Cronbach alpha
1. Breathing	0.33	0.28	0.81	0.31	0.26	0.80	0.29	0.25	0.83
2. Food	0.59	0.50	0.80	0.57	0.50	0.79	0.57	0.52	0.81
3. Rest	0.64	0.56	0.80	0.63	0.55	0.78	0.63	0.56	0.81
4. Sleep	0.64	0.54	0.80	0.62	0.52	0.79	0.67	0.59	0.80
5. Hygiene	0.48	0.37	0.81	0.52	0.42	0.79	0.55	0.47	0.81
6. Communication	0.34	0.31	0.81	0.37	0.35	0.80	0.29	0.27	0.83
7. Support	0.18	0.11	0.82	0.18	0.09	0.81	0.17	0.07	0.83
8. Return to work	0.62	0.48	0.80	0.64	0.49	0.79	0.67	0.53	0.81
9. Feeling in control	0.71	0.64	0.79	0.69	0.61	0.78	0.80	0.75	0.80
10. Well-being	0.70	0.63	0.79	0.68	0.61	0.78	0.77	0.71	0.80
11. Moderate pain	0.50	0.34	0.81	0.54	0.37	0.81	0.50	0.31	0.84
12. Severe pain	0.64	0.53	0.80	0.60	0.48	0.79	0.64	0.54	0.81
13. Nausea/Vomiting	0.41	0.30	0.81	0.41	0.32	0.80	0.31	0.23	0.83
14. Anxiety	0.60	0.51	0.80	0.56	0.46	0.79	0.69	0.63	0.80
15. Depressed	0.52	0.44	0.80	0.53	0.45	0.79	0.70	0.63	0.80
Test scale			**0.81**			**0.80**			**0.83**

non-responders.[14] This highlights the QoR-15's clinical usefulness, not only for patients, but also for staff using the QoR-15 for research and quality improvement purposes.[14]

The QoR-15's brevity means it can be read and completed quickly, as opposed to other longer QoR scores.[8 15 16] Currently, one of the most well-regarded and widely used QoR scores in surgery is the QoR-40.[4] Myles and colleagues[6] developed and psychometrically evaluated this comprehensive 40-item score. However it is a lengthy questionnaire, with most patients taking around 10 min to complete it. By contrast, in our study, the measured subset of patients was able to complete the QoR-15 questionnaire in an average of 2.6 min. This is slightly longer than the original development and validation paper,[7] but this could reflect that the QoR-15 questionnaire in our study was not self-administered.

Most patients' QoR-15 scores had returned to their preoperative values by 48 h and exceeded them by seven days. This indicates the preoperative score may not be a true baseline score. Focusing in on individual items of the score, the results indicate that patients are tired, anxious and in pain in the 24 h before surgery. These circumstances may not provide an ideal baseline for comparison. A measurement taken during preoperative assessment, or at the time of surgical booking could provide a truer baseline score, with possibly lower scores for anxiety and tiredness. However, preoperative pain measures may be unchanged, as pain may be the reason for the surgery. A QoR-15 measurement after complete recovery may be a better comparator, but this assumes that patients will have a complete recovery.

Measuring the QoR-15 at three postoperative time points is feasible, but is very time and resource heavy, requiring dedicated staff. The data suggest most patients return to their preoperative scores by 48 h after surgery, despite this score not being a true baseline. If a patient has not approximated their preoperative score by 48 h, this may indicate a deviation from their expected recovery. Enhanced recovery after surgery is based on adherence to protocols and the care pathway, and managing deviations appropriately.[16] Measurement of the QoR-15 at 48 h may aid identification of patients who are not recovering as well as expected, and allow targeted intelligent interventions to aid their recovery. We suggest one follow-up call at 48 h would enable an adequate patient-centred assessment of postoperative recovery after day case orthopaedic surgery.

The original paper assessed 21 day case surgery patients who were contacted by telephone the day after their surgery.[7] Their mean QoR-15 scores and kurtosis were consistent with a normal distribution. This is opposed to our analysis where we saw an increasingly negative skew to the data, reflecting either better patient recovery or lower surgical severity. Floor or ceiling effects are present if greater than 15% of subjects achieve the highest or lowest possible scores.[17] This was not seen in the preoperative, 24 h or 48 h scores, however a ceiling effect was observed seven days postoperatively, with 17% of patients achieving the highest score. This is an expected effect, as patients will hopefully continue to recover from their surgery over time. In the longer term, the increasing percentage of patients achieving the highest possible score is a patient-reported outcome measure of surgical success.

The QoR-15 demonstrated strong construct validity. It was able to discriminate between the genders, as it has previously been shown that women have a worse postoperative recovery.[2 18] A negative association was demonstrated between the QoR-15 and duration of surgery and total opioid use. The negative association with total opioid use may reflect the severity of the surgery, as a bigger, more painful operation may lead to a slower recovery post-discharge.

Internal consistency was measured using Cronbach's α coefficient. The results were high and satisfied published recommendations (0.70–0.90).[19] These results are comparable with the validation paper of the QoR-15,[7] the longer form QoR-40[6] and exceed those of the modified 28-item QoR 40.[12] Internal consistency was also measured using inter-item correlation. Each item was internally consistent (coefficient values 0.79–0.83) and correlated well with the total QoR-15 score.

The responsiveness was assessed using Cohen effect size and standardized response means.[20] The Cohen effect size was 0.37, suggesting a moderate ability to detect change. This is a lower

Table 4 Responsiveness of the total QoR-15 score. Individual item mean QoR-15 scores, Cohen effect sizes and Standardized Response Means (SRM) at 24 h, 48 h and seven days after surgery, compared with preoperative values (437 patients)

	Mean scores (±SD)	Mean change from pre-op score (95% CI)	% Change from pre-op	Cohen effect size	SRM
QoR-15 score pre-op					
1. Breathing	9.6 (1.0)	NA	NA	NA	NA
2. Food	9.6 (1.2)	NA	NA	NA	NA
3. Rest	8.2 (2.1)	NA	NA	NA	NA
4. Sleep	7.8 (2.2)	NA	NA	NA	NA
5. Hygiene	9.7 (1.1)	NA	NA	NA	NA
6. Communication	9.9 (0.6)	NA	NA	NA	NA
7. Support	9.4 (1.7)	NA	NA	NA	NA
8. Return to work	8.5 (2.7)	NA	NA	NA	NA
9. Feeling in control	8.8 (2.0)	NA	NA	NA	NA
10. Well-being	8.6 (2.0)	NA	NA	NA	NA
11. Moderate pain	6.6 (3.4)	NA	NA	NA	NA
12. Severe pain	8.1 (3.0)	NA	NA	NA	NA
13. Nausea/vomiting	9.4 (2.1)	NA	NA	NA	NA
14. Anxiety	7.6 (2.7)	NA	NA	NA	NA
15. Depressed	8.6 (2.5)	NA	NA	NA	NA
QoR-15 Score at 24 H					
1. Breathing	9.8 (0.9)	0.2 (0.05 to 0.3)	2.0	+0.2	+0.1
2. Food	9.1 (2.1)	−0.6 (−0.8 to −0.4)	6.3	−0.5	−0.3
3. Rest	8.4 (2.4)	0.2 (−0.08 to 0.5)	2.4	+0.1	+0.07
4. Sleep	7.7 (2.8)	−0.1 (−0.4 to 0.2)	1.3	−0.05	−0.03
5. Hygiene	8.8 (2.3)	−0.9 (−1.1 to −0.7)	9.3	−0.8	−0.4
6. Communication	9.9 (0.5)	0.02 (−0.05 to 0.1)	0.2	+0.03	+0.03
7. Support	9.8 (1.2)	0.4 (0.2 to 0.6)	4.3	+0.2	+0.2
8. Return to work	5.2 (3.6)	−3.2 (−3.6 to −2.8)	37.6	−1.2	−0.8
9. Feeling in control	8.5 (2.3)	−0.3 (−0.5 to −0.01)	3.4	−0.2	−0.1
10. Well-being	8.8 (2.1)	0.2 (−0.06 to 0.4)	2.3	+0.1	+0.07
11. Moderate pain	6.2 (3.4)	−0.4 (−0.9 to −0.002)	6.1	−0.1	−0.09
12. Severe pain	8.6 (2.9)	0.5 (0.1 to 0.8)	6.2	+0.2	+0.1
13. Nausea/vomiting	9.1 (2.3)	−0.2 (−0.5 to 0.05)	2.1	−0.1	−0.08
14. Anxiety	9.0 (2.2)	1.3 (1.1 to 1.6)	17.1	+0.5	+0.4
15. Depressed	9.3 (1.9)	0.7 (0.4 to 0.9)	8.1	+0.3	+0.3
QoR-15 Score at 48 H					
1. Breathing	9.9 (0.8)	0.2 (0.1 to 0.3)	2.1	+0.2	+0.2
2. Food	9.3 (1.7)	−0.3 (−0.5 to −0.1)	3.1	−0.3	−0.2
3. Rest	8.8 (2.0)	0.6 (0.3 to 0.8)	7.3	+0.3	+0.2
4. Sleep	8.5 (2.3)	0.6 (0.4 to 0.9)	7.7	+0.3	+0.2
5. Hygiene	9.0 (1.9)	−0.7 (−0.9 to −0.5)	7.2	−0.6	−0.4
6. Communication	9.9 (0.4)	0.04 (−0.03 to 0.1)	0.4	+0.07	+0.05
7. Support	9.7 (1.4)	0.3 (0.09 to 0.5)	3.2	+0.2	+0.1
8. Return to work	5.8 (3.4)	−2.7 (−3.0 to −2.3)	31.8	−1.0	−0.7
9. Feeling in control	8.8 (2.1)	0.005 (−0.2 to 0.2)	0.06	+0.003	+0.002
10. Well-being	9.0 (1.9)	0.3 (0.1 to 0.5)	3.5	+0.2	+0.2
11. Moderate pain	6.3 (3.3)	−0.3 (−0.7 to 0.1)	4.5	−0.09	−0.07
12. Severe pain	8.6 (2.7)	0.5 (0.2 to 0.9)	6.2	+0.2	+0.1
13. Nausea/vomiting	9.5 (1.5)	0.2 (−0.06 to 0.4)	2.1	+0.1	+0.07
14. Anxiety	9.1 (2.0)	1.5 (1.2 to 1.8)	19.7	+0.6	+0.5
15. Depressed	9.4 (1.8)	0.8 (0.5 to 1.0)	9.3	+0.3	+0.3
QoR-15 Score at seven days					
1. Breathing	9.9 (0.6)	0.3 (0.2 to 0.4)	3.1	+0.3	+0.3
2. Food	9.7 (1.1)	0.08 (−0.06 to 0.2)	0.8	+0.07	+0.05
3. Rest	9.4 (1.5)	1.1 (0.9 to 1.3)	13.4	+0.5	+0.5
4. Sleep	9.1 (1.9)	1.2 (1.0 to 1.4)	15.4	+0.5	+0.5
5. Hygiene	9.4 (1.6)	−0.3 (−0.4 to −0.1)	3.1	−0.3	−0.2
6. Communication	10.0 (0.3)	0.07 (0.02 to 0.1)	0.7	+0.1	+0.1
7. Support	9.7 (1.4)	0.3 (0.1 to 0.5)	3.2	+0.2	+0.1
8. Return to work	7.3 (3.0)	−1.2 (−1.5 to −0.9)	14.1	−0.4	−0.3
9. Feeling in control	9.4 (1.5)	0.6 (0.4 to 0.8)	6.8	+0.3	+0.3

Continued

Table 4 *Continued*

	Mean scores (±SD)	Mean change from pre-op score (95% CI)	% Change from pre-op	Cohen effect size	SRM
10. Well-being	9.3 (1.6)	0.7 (0.5 to 0.9)	8.0	+0.4	+0.4
11. Moderate pain	7.5 (3.0)	0.9 (0.5 to 1.3)	13.6	+0.3	+0.2
12. Severe pain	9.2 (2.1)	1.1 (0.8 to 1.4)	13.6	+0.4	+0.3
13. Nausea/vomiting	9.8 (1.1)	0.4 (0.2 to 0.6)	4.3	+0.2	+0.2
14. Anxiety	9.5 (1.5)	1.8 (1.6 to 2.1)	23.7	+0.7	+0.7
15. Depressed	9.5 (1.6)	0.8 (0.6 to 1.1)	9.3	+0.3	+0.3

value than the original validation paper, but this could reflect the lower severity of procedures undertaken in our ambulatory surgery population. For individual dimensions, the effect sizes varied, being greatest for the physical independence questions, and lowest for the pain dimension. For all individual items, Cohen effect size values ranged from small to large scales of responsiveness. The most responsive was question eight, on the ability to return to work or usual activities, followed by question five, on ability to maintain personal toilet and hygiene unaided. The least responsive items varied between the three time points, and included the items on 'feeling comfortable and in control' and 'the ability to communicate with family and friends'. This raises the possibility of removing these two questions from the QoR-15 score for day case surgery. However, they address important discrete parts of a patient's recovery and wellbeing, and patients scoring low in these items must be identified.

Our study has some limitations. It was conducted in a single university-affiliated hospital in London, UK, therefore generalizability outside this setting is unknown; however, our patient cohort was representative of diverse ethnicity, age, gender and comorbidities. The study cohort was limited to orthopaedic day case surgery, therefore, formal assessment in other procedural cohorts may be of interest. Finally, the follow-up questionnaires were not self-administered, which may have led to administration bias. To address some of these limitations, we suggest that future work may focus on multi-centre studies assessing patients undergoing other types of surgery, and validation of an app-based QoR-15 score for patient self-administration,

In conclusion, the QoR-15 is a valid, responsive, reliable, acceptable and feasible tool for measuring the quality of a patient's postoperative recovery after day case orthopaedic surgery. We propose that measurement of QoR-15 before surgery (but not on day of surgery) and 48 h postoperatively should provide a useful and feasible assessment of patient-reported outcome after day case orthopaedic surgery, which may be applied both in clinical studies and for assessing the impact of changes in healthcare delivery.

Authors' contributions

Study design/planning: M.C., D.K., S.R.M.
Study conduct: M.C., M.A.R.
Data analysis: E.M.K.W., S.R.M.
Writing paper: M.C., E.M.K.W., S.R.M.
Revising paper: all authors

Supplementary material

Supplementary material is available at *British Journal of Anaesthesia* online.

Acknowledgements

We would like to thank the following contributors who collected data: R. Aziz., E.J. Bettini., D. Blackwood., D. Brunnen., L. Cairns., R. Coe., R. Gordon-Williams., P. Gorur., M. Hoy., E.I. Mcllroy., J. Patel., D. Wagstaff., D. Wyndham., D. Zeloof.

Declaration of interest

None declared.

Funding

National Institute for Health Research UCL/UCLH Biomedical Research Centre (where S.R.M. is a member of the Faculty), through support for the UCL/UCLH Surgical Outcomes Research Centre, University College Hospital, London. S.R.M. receives funding for her role as Deputy Director of the NIAA Health Services Research Centre and as a Health Foundation Improvement Science Fellow (2015-18). M.C. was supported through a grant awarded by the National Institute for Academic Anaesthesia's Health Services Research Centre.

References

1. Ng L, Mercer-Jones M. Day case surgery guidelines. *Surg* 2014; **32**: 73–8
2. Hall JA, Dornan MC. Patient sociodemographic characteristics as predictors of satisfaction with medical care: A meta-analysis. *Soc Sci Med* 1990; **30**: 811–8
3. Aldrete JA, Kroulik D. A Postanesthetic Recovery Score. *Anesth Analg* 1970; **49**: 924–34
4. Herrera FJ, Wong J, Chung F. A systematic review of postoperative recovery outcomes measurements after ambulatory surgery. *Anesth Analg* 2007; **105**: 63–9
5. Talamini MA, Stanfield CL, Chang DC, Wu AW. The surgical recovery index. *Surg Endosc Other Interv Tech* 2004; **18**: 596–600
6. Myles PS, Weitkamp B, Jones K, Melick J, Hensen S. Validity and reliability of a postoperative quality of recovery score: the QoR-40. *Br J Anaesth* 2000; **84**: 11–5
7. Stark PA, Myles PS, Burke JA. Development and Psychometric Evaluation of a Postoperative Quality of Recovery Score: The QoR-15. *Anesthesiology* 2013; **118**: 1332–40
8. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis* 1985; **38**: 27–36
9. Ware JE Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. *Med Care* 1996; **34**: 220–33
10. Kazis L, Anderson J, Meenan R. Effect sizes for interpreting changes in health status. *Med Care* 1989; **27**: 178–9

11. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; **16**: 297–334
12. Idvall E, Berg K, Unosson M, Brudin L, Nilsson U. Assessment of recovery after day surgery using a modified version of quality of recovery-40. *Acta Anaesthesiol Scand* 2009; **53**: 673–7
13. Bost J, Williams B, Bottegal M, Dang Q, Rubio D. The 8-Item Short-Form Health Survey and the Physical Comfort Composite Score of the Quality of Recovery 40-Item Scale Provide the Most Responsive Assessments of Pain, Physical Function, and Mental Function During the First 4 Days After Ambulatory Knee Surgery with Regional Anesthesia. *Anesth Analge* 2007; **105**: 1693–700
14. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998; **2**: 1–74
15. Royse CF, Newman SD, Chung F, *et al*. Development and Feasibility of a Scale to Assess Postoperative Recovery: The Post-operative Quality Recovery Scale. *Anesthesiology* 2010; **113**: 892–905
16. Fearon KCH, Ljungqvist O, Meyenfeldt MV, *et al*. Enhanced recovery after surgery: A consensus review of clinical care for patients undergoing colonic resection. *Clin Nutr* 2005; **24**: 466–77
17. Terwee CB, Bot SDM, De boer MR, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; **60**: 34–42
18. Blanchard CG, Labrecque MS, Ruckdeschel JC, Blanchard EB. Physician behaviors, patient perceptions, and patient characteristics as predictors of satisfaction of hospitalized adult cancer patients. *Cancer* 1990; **65**: 186–92
19. Mcdowell I. Measuring Health : A Guide to Rating Scales - third edition. 2006
20. Norman GR, Wyrwich K, Patrick D. The mathematical relationship among different forms of responsiveness coefficients. *Qual Life Res* 2007; **16**: 815–22

Handling editor: J. P. Thompson