International Journal of Obstetric Anesthesia (2019) 39, 51–59
0959-289X/$ - see front matter
https://doi.org/10.1016/j.ijoa.2019.01.010

ELSEVIER
www.obstetanesthesia.com

ORIGINAL ARTICLE

Evaluation of the Obstetric Quality-of-Recovery score (ObsQoR-11) following non-elective caesarean delivery

S. Ciechanowicz,[a] R. Howle,[a] C. Heppolette,[a] B. Nakhjavani,[a] B. Carvalho,[b] P. Sultan[c]
[a]*University College London Hospital, London, UK*
[b]*Stanford University School of Medicine, Stanford, CA, USA*
[c]*University College London Hospital, Honorary Senior Lecturer, University College London, London, UK*

ABSTRACT
Background: Few robust scoring tools exist to assess recovery following caesarean delivery (CD). We evaluated a new obstetric quality of recovery score (ObsQoR-11, initially formulated for elective CD) following non-elective CD.
Methods: ObsQoR-11 questionnaires were completed by women at day one post non-elective CD. Convergent validity was assessed by correlation of ObsQoR-11 with a 100 mm numerical rating scale (NRS) of general health status; discriminant validity by correlation with good vs poor recovery (NRS of $\geq$70 vs <70 mm, respectively); and content validity by correlation with length of stay (LOS), CD category, parity, gestation, previous CD, duration, blood loss, haemoglobin, age and body mass index. Cronbach's alpha, inter-item, split-half and test–retest correlation assessed reliability. Feasibility was tested by recruitment rate and time for ObsQoR-11 completion.
Results: One hundred women completed ObsQoR-11 at 24 h and 20 women repeated it at 25 h. ObsQoR-11 correlated strongly with NRS (r = 0.72 [95% CI 0.61 to 0.81], P <0.0001); discriminated well between good versus poor recovery (median [IQR] score 97 [86.5–101] vs 64 [50.5–78.5], P <0.0001); correlated to LOS (r = −0.24 [−0.42 to −0.04], P=0.02) and parity (r = 0.24 [0.04 to 0.42], P=0.02). Reliability was acceptable: Cronbach's alpha 0.75; inter-item correlation >0.15; split-half reliability 0.96; and intra-class correlation >0.6; no floor or ceiling effects. One hundred percent completed the ObsQoR-11 (median [IQR] completion time 117 [89–156] s).
Conclusions: ObsQoR-11 is valid and reliable in assessing recovery after non-elective CD. Further research should assess generalisability and use following vaginal delivery.

Keywords: Recovery; Quality; Caesarean; Postoperative

Introduction

Caesarean delivery (CD) is one of the most frequently performed surgical procedures worldwide. However, quality of recovery scoring tools for this patient population are limited. Recovery after surgery is a multi-dimensional process but many prospective studies evaluating post-partum recovery use single-dimensional outcome measures, such as visual analogue scale pain scores.[1–3] These outcome measures, although quick and easy to assess, do not provide a holistic indication of patient-reported functional recovery.[4] The QoR-40 (at 24 h),[5] QoR-15 (at 12 and 24 h),[6] EuroQol-5D (at 24–48 h)[7] and Short Form Health Survey (SF-36 at six weeks and six months),[8] are examples of multi-dimensional recovery scoring tools that have been used following CD. Although they can evaluate postoperative recovery by measuring key patient-reported recovery domains,[9,10] none was developed or validated in obstetric patients following CD. The EuroQol-5D has been used in this context at 24 h[7,11] and at 12 months,[12] but has not been validated in the obstetric setting. These scoring tools also neglect items such as the ability to care for the newborn child, which is indicative of good recovery and necessary to permit discharge from the postnatal ward. Finally, recovery tools which include assessment of recovery from anaesthesia utilise items specific to recovery from general, rather than neuraxial, anaesthesia (for example postoperative sore throat).

There is a need to develop, evaluate and validate a simple multi-dimensional obstetric recovery assessment

Accepted January 2019
Correspondence to: Dr S. Ciechanowicz, Department of Anaesthesia and Perioperative Medicine, University College London, 235 Euston Rd., London, NW1 2BU, United Kingdom.
E-mail address: s.ciechanowicz@nhs.net

tool for use following CD, to quantify standards of postoperative recovery. The Obstetric Quality-of-Recovery Score (ObsQoR-11) has recently been formulated to evaluate recovery in the first 24 h following elective CD.[13] Patient-reported recovery items were initially selected from the QoR-40 scoring tool,[14] and following expert evaluation obstetric-specific items were added. Stakeholders subsequently selected and ranked items considered of greatest importance for testing in women undergoing elective CD. This identified the strongest performing recovery items and permitted evaluation of the scoring tool through its correlation with other measures of recovery such as global health status and length of hospital stay (LOS). The development study on elective CD patients suggests that ObsQoR-11 is a promising tool for elective surgery: ObsQoR-11 correlated with global health status at 24 h post CD, was negatively correlated to LOS and performed well in measures of reliability and feasibility.[13]

The ObsQoR-11 may be used to aid early identification of patients at risk of prolonged hospital stay. Clinically, ObsQoR-11 scores may assist clinical decision-making regarding requirement for senior review, aid evaluation of functional recovery, readiness for discharge on day one post-partum as part of enhanced recovery programmes and may be used as a quality improvement tool, which may help improve patient experience following obstetric surgery.

Validation of the ObsQoR-11 following non-elective CD could increase its scope for widespread adoption. This study aimed to evaluate the ObsQoR-11 following non-elective CD by testing validity, reliability, clinical acceptability and feasibility in this setting.

Methods

A prospective observational study of term women undergoing non-elective CD (Royal College of Obstetricians and Gynaecologist's classification of urgency category 1, 2 and 3),[15] was conducted between November 2017 and March 2018 at University College London Hospital, UK. The local research office designated the study as service evaluation and it was therefore exempt from research ethics committee approval, consistent with a previous study performed at our institution.[9] The ObsQoR-11 questionnaire (with the same questions as previously published,[13] but using a modified layout with additional stylised faces to make it simpler to understand based on patient feedback) (Fig. 1), was completed by women 24 h after the documented end of surgery (and 20% of these women also completed the questionnaire at 25 h).

Women were included if they underwent non-elective CD under neuraxial anaesthesia at $\geq$37 weeks' gestational age. Women were excluded if they were less than 18 years-of-age, refused to participate or were unable to read or understand written English.

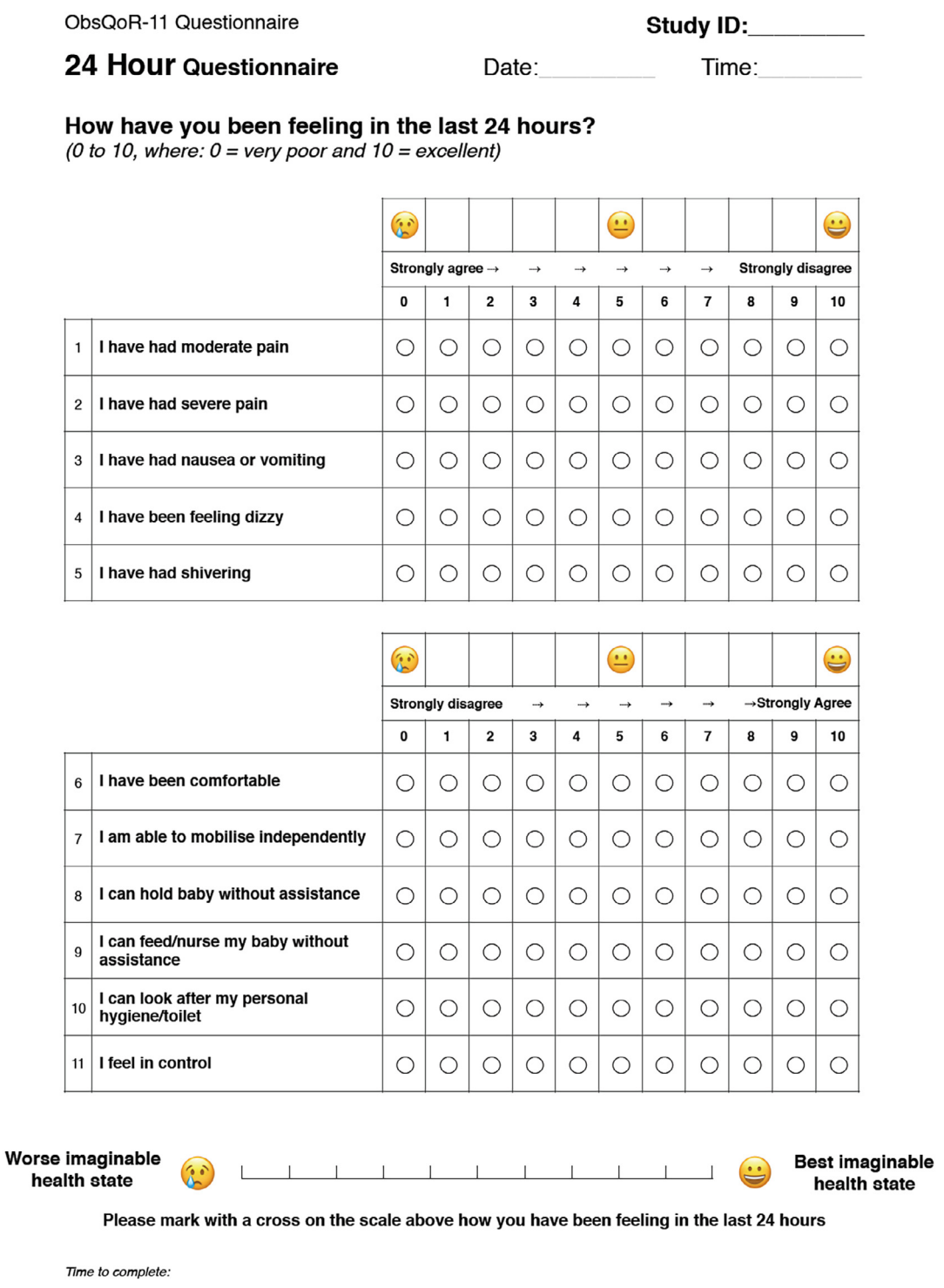
Over the study period, recruitment days corresponded to investigator availability. No patient was awoken during the night and so recruitment was limited to non-elective CDs performed during daytime working hours (08:00–18:00); outside of these hours, women were indirectly excluded. Three investigators interviewed the women: (R.H, C.H and B.N), none was involved in clinical care. Baseline demographic and clinical data were collected at the time of enrolment for descriptive purposes. Women were approached whilst on the postnatal ward at 24 ($\pm$2) h after surgery and asked to complete the ObsQoR-11 questionnaire.

The questionnaire asked women to rate each recovery item with an 11-point numerical Likert scale (0 = strongly negative; 10 = strongly positive; Fig. 1). During the interview, women were also asked to measure their general health status using a global health numerical rating scale (NRS) represented by a 100 mm line marked at each end with anchors, 'worst imaginable health state' to 'best imaginable health state' and with 'sad' or 'happy' stylised representations of faces (Fig. 1). The investigator confirmed patient understanding of these statements. A subset of women were also asked to repeat the ObsQoR-11 questionnaire and global health NRS at 25 h postoperatively to test reliability;[10] these women were randomly chosen by the investigator.

Our institutional neuraxial anaesthesia regimen includes intrathecal administration of hyperbaric bupivacaine 12–14 mg with diamorphine 300 μg via either a single-shot spinal or combined spinal-epidural technique. If an epidural top-up was performed, up to 20 mL lignocaine 2% with adrenaline 1:200 000 and fentanyl 100 μg or diamorphine 3 mg was administered. Our standard anaesthetic protocol was followed for all women who completed the ObsQoR-11 scoring tool. For postoperative analgesia, women received regular paracetamol 1 g and ibuprofen 400 mg four times daily unless contraindicated, and for breakthrough pain as required oral morphine 20 mg four-hourly on the day of surgery and dihydrocodeine 30 mg six-hourly from day one postoperatively. Intravenous ondansetron 4–8 mg and cyclizine 50 mg eight-hourly as required were also prescribed unless contraindicated.

The primary aim of the study was to test the validity of the ObsQoR-11 following non-elective CD. This was achieved by the following:

(a) Validity as a measure of accuracy. The ObsQoR-11 was assessed by two subtypes of construct validity: convergent and discriminant validity; and by content validity, as outlined. For convergent validity the ObsQoR-11 scores at 24 and 25 h were correlated to 100-mm NRS assessment of global health status at the same times. For discriminant validity a comparison was made between the 24 h ObsQoR-11 scores of women who had a 'good' or 'poor' postoperative recovery, defined by global NRS assessment scores of


ObsQoR-11 Questionnaire

Study ID:_____

24 Hour Questionnaire Date:_____ Time:_____

How have you been feeling in the last 24 hours?

(0 to 10, where: 0 = very poor and 10 = excellent)

		😢					😐					😀
		Strongly agree →		→	→	→	→	→	→			Strongly disagree
		0	1	2	3	4	5	6	7	8	9	10
1	I have had moderate pain	○	○	○	○	○	○	○	○	○	○	○
2	I have had severe pain	○	○	○	○	○	○	○	○	○	○	○
3	I have had nausea or vomiting	○	○	○	○	○	○	○	○	○	○	○
4	I have been feeling dizzy	○	○	○	○	○	○	○	○	○	○	○
5	I have had shivering	○	○	○	○	○	○	○	○	○	○	○

		😢					😐					😀
		Strongly disagree		→	→	→	→	→	→			→Strongly Agree
		0	1	2	3	4	5	6	7	8	9	10
6	I have been comfortable	○	○	○	○	○	○	○	○	○	○	○
7	I am able to mobilise independently	○	○	○	○	○	○	○	○	○	○	○
8	I can hold baby without assistance	○	○	○	○	○	○	○	○	○	○	○
9	I can feed/nurse my baby without assistance	○	○	○	○	○	○	○	○	○	○	○
10	I can look after my personal hygiene/toilet	○	○	○	○	○	○	○	○	○	○	○
11	I feel in control	○	○	○	○	○	○	○	○	○	○	○

Worse imaginable health state 😢 😐 **Best imaginable health state**

Please mark with a cross on the scale above how you have been feeling in the last 24 hours

Time to complete:

Fig. 1 Modified ObsQoR-11 survey instrument and Global Health NRS. NRS: numerical rating scale (revised by Dr Derek Brunnen)

≥70 versus <70 mm at 24 h, respectively. For content validity the correlation between ObsQoR-11 score and LOS (number of nights), age (<30 years versus >35 years), body mass index (BMI), parity, gestational age, gestation (singleton versus multiple), previous CD, urgency of surgery (category 1–3), duration of surgery, estimated blood loss during surgery (<500 mL versus >1L), pre- and post-operative haemoglobin concentration, and change in haemoglobin concentration (preoperative to postoperative day one was analysed). These outcomes were selected on the basis of previous studies demonstrating a relationship and bio-

logical plausibility. Given the well-established increase in peri-operative risk with advancing maternal age and high BMI,[16,17] we hypothesised that these could negatively correlate with quality of recovery. Women with a previous delivery might report a better ObsQoR-11 score due to experience with the post-partum period. We also hypothesised that women with a multiple gestation pregnancy would have a poorer ObsQoR-11 score due to a higher incidence of complications and greater requirement for nursing and childcare compared with singleton pregnancy.[16] In addition, we predicted poorer recovery at 24 h associated with an extended duration of surgery and greater blood loss or change in haemoglobin concentration, as demonstrated in the obstetric and non-obstetric literature.[18,19]

(b) Reliability as a measure of consistency of the ObsQoR-11 (which refers to the agreement between multiple items, with internal consistency gauging how well a test measures what it is intended to measure) was assessed by evaluating: (1) Internal consistency using Cronbach's alpha, inter-item correlation tests and split-half reliability (evaluation of the correlation between random split segments of ObsQoR-11 items). (2) The test-retest reliability in a subset of women who were asked to repeat the questionnaire 30–60 min later (at 25 h), allowing correlation to 24 h responses. (3) The floor and ceiling effects by evaluating whether <15% respondents achieved either the highest (110) or lowest possible score (0).

(c) Acceptability and feasibility were assessed by evaluating: (1) Recruitment rate as determined by the percentage of women agreeing to complete the scoring tool; (2) The successful completion rate, by recording the number of correctly completed ObsQoR-11 forms, without missing data; (3) The time taken to complete the ObsQoR-11 tool, measured using a stopwatch and recorded by the investigator. (this was not disclosed to the patient.)

The sample size calculation was guided by previous studies, as power calculation is not reliable for correlation analysis.[10] Data are presented as mean (SD), median [interquartile range (IQR)], number (%), and 95% confidence intervals (CI) as appropriate. Continuous data were tested for normality using the Shapiro-Wilk normality test; all percentages were rounded up to the nearest integer. Correlations between ObsQoR-11 items and global health NRS scores were determined using Spearman rank correlation coefficients (r). Non-parametric data were compared using the Wilcoxon signed-rank test. Internal consistency was measured with Cronbach's alpha[20] and split-half reliability. Test–retest reliability was measured by intra-class correlation coefficient (ICC). The Kruskal-Wallis test was used as a non-parametric method to compare modes of anaesthesia. Statistical analysis was performed using GraphPad Prism (7.0, U.S.A). The null hypothesis was rejected if the two-tailed *P*-value was <0.05.

Results

During the study period, 100 women were recruited. All completed the ObsQoR-11 questionnaire at 24 h postoperatively and 20 of these women repeated it at 25 h. There were no incomplete or missing data and no exclusions from analysis. Clinical characteristics of the women and their CDs are presented in Table 1.

Table 1 Summary of demographic and clinical characteristics of women completing the ObsQoR-11 questionnaire

Age (y)	
Mean ± SD	33.1 ± 5.0
Range	18–44
Body mass index	
Median (IQR)	23.3 (21.5–26.5)
Range	17–42
Parity	
0	72
1	19
2	5
3	2
≥4	2
Gestation	
Single	97
Multiple	3
Length of hospital stay (nights)	
Median (IQR)	3 (2–4)
Range	1–8
Pre-existing medical conditions (%)	
Respiratory	12
Cardiovascular	0
Neurological	2
Endocrine	16
Haematological	3
Musculoskeletal	3
Psychiatric	4
Other	2
Category of emergency CD (%)	
1	19
2	51
3	30
Obstetric indication for CD (%)	
Pathological CTG	35
Failure to progress	25
Previous CD	17
Breech	13
Failed IOL	8
Pre-eclampsia	1
Uncontrolled DM	1
Previous CD (%)	
Yes	17
No	83
Anaesthesia technique (%)	
Epidural top-up	40
Single-shot spinal	55
Combined spinal-epidural	5

n=100; IQR: interquartile range; SD: standard deviation; CD: caesarean delivery; CTG: cardiotocogaph; IOL: induction of labour; DM: diabetes mellitus

ObsQoR-11 scores at 24 h correlated with global health NRS scores at 24 h (Spearman r = 0.72 (95% CI 0.61 to 0.81); $P < 0.0001$). Individual item correlation to global health score is demonstrated in Table 2. The ObsQoR-11 scores differed significantly for women who had a 'good' or 'poor' postoperative recovery. ObsQoR-11 data were not normally distributed. Median [IQR] ObsQoR-11 scores at 24 h were 97 [86.5–101] for 'good' recovery versus 64 [50.5–78.5] for 'poor' recovery ($P < 0.0001$).[10]

Median [IQR] LOS for women in the study was three nights[2–4] with a range of one to eight nights (Table 1). The ObsQoR-11 scores at 24 h had a weak negative correlation with LOS, (r = −0.24; CI −0.42 to −0.04; P=0.02), suggesting that higher scores were associated with a reduced LOS. There was a weak positive correlation of scores with parity (r = 0.24; CI 0.04 to 0.42; P=0.02), but no significant correlation with other clinical characteristics (Table 3). There were no differences in scores for women with blood loss ≤500 mL versus blood loss ≥1 L (median 78 vs 71.5, P=0.51), or maternal age <30 years vs >35 years (median 79 vs. 77, P=0.59). There was no significant difference in the median scores for modality of anaesthesia used (epidural top-up, spinal or combined spinal-epidural: 72.5, 83, 92 respectively, P=0.10), or for category of non-elective CD (1, 2 or 3: 69.5, 75.5, 86.5 respectively; P=0.09).

Internal consistency measured using Cronbach's alpha was 0.75. The inter-item correlation matrix for ObsQoR-11 is outlined in Table 4. Inter-item correlations were mostly r > 0.15, a good indicator of consistency. Split-half reliability with Spearman Brown adjustment (which measures the extent to which all parts of the test contribute equally to the desired measurement) was 0.96, implying an equal contribution from all items. Test–retest reliability (represented by intraclass correlation coefficient) of the ObsQoR-11 score demonstrated an r_i of ≥0.62 (range 0.62–0.98) for all

Table 2 Summary of correlations of ObsQoR-11 items to global health numerical rating scale

ObsQoR-11 Item	Correlation to global health NRS score* Spearman r (95% CI)	*P*-value
Moderate pain	0.32 (0.12 to 0.49)	0.0014
Severe pain	0.51 (0.34 to 0.65)	<0.0001
Nausea or vomiting	0.41 (0.22 to 0.56)	<0.0001
Dizzy	0.35 (0.16 to 0.52)	0.0003
Shivering	0.42 (0.24 to 0.57)	<0.0001
Comfortable	0.54 (0.38 to 0.67)	<0.0001
Mobilise independently	0.54 (0.38 to 0.67)	<0.0001
Able to hold baby	0.41 (0.23 to 0.56)	<0.0001
Able to nurse/feed baby	0.44 (0.26 to 0.59)	<0.0001
Able to take care of personal hygiene	0.53 (0.36 to 0.66)	<0.0001
Feeling in control	0.58 (0.44 to 0.71)	<0.0001

*Global health numerical rating scale (NRS), represented as a 100 mm line, marked at each end with anchors 'worst imaginable health state'; to 'best imaginable health state' and with 'sad' or 'happy' stylised representations of faces.

Table 3 Summary of correlations of clinical characteristics to ObsQoR-11 score

Clinical characteristic	Correlation to ObsQoR-11 Spearman r (95% CI)	*P*-value
LOS	−0.24 (−0.42 to −0.04)	0.02
Parity	0.24 (0.04 to 0.42)	0.02
Gestational age	−0.08 (−0.28 to 0.12)	0.42
Gestation (singleton vs twins)	0.11 (−0.9 to 0.31)	0.27
Maternal age	0.07 (−0.13 to 0.27)	0.46
BMI	−0.11 (−0.32 to 0.10)	0.28
Category of CD	0.14 (−0.07 to 0.23)	0.78
Previous CD	0.08 (−0.12 to 0.28)	0.43
Duration of surgery	−0.02 (−0.10 to 0.31)	0.28
Blood loss	−0.08 (−0.28 to 0.12)	0.42
Pre Hb	0.14 (−0.07 to 0.33)	0.17
Post Hb	0.08 (−0.13 to 0.28)	0.45
Change in Hb	−0.09 (−0.29 to 0.12)	0.39

LOS: length of hospital stay (number of nights); BMI: body mass index; CD: caesarean delivery; Pre Hb: pre-operative haemoglobin concentration; Post Hb: postoperative haemoglobin concentration; Change in Hb: difference from pre- to postoperative haemoglobin concentration.

Table 4 Inter-item correlation matrix for ObsQoR-11 following non-elective caesarean delivery

ObsQoR-11 item number	Global health NRS*	Total ObsQoR-11 score	1	2	3	4	5	6	7	8	9	10	11
1	0.32	0.35	–										
2	0.51	0.68	0.28	–									
3	0.41	0.56	0.22	0.31	–								
4	0.35	0.52	0.19	0.17	0.47	–							
5	0.42	0.61	0.09	0.47	0.36	0.46	–						
6	0.54	0.60	0.11	0.46	0.32	0.17	0.32	–					
7	0.54	0.71	0.18	0.48	0.37	0.30	0.35	0.49	–				
8	0.41	0.64	0.05	0.29	0.19	0.23	0.18	0.30	0.46	–			
9	0.44	0.63	0.11	0.25	0.13	0.18	0.17	0.26	0.42	0.72	–		
10	0.53	0.73	0.11	0.41	0.34	0.27	0.22	0.44	0.62	0.57	0.58	–	
11	0.59	0.75	0.23	0.50	0.32	0.19	0.36	0.48	0.53	0.48	0.55	0.74	–

Obstetric Quality of Recovery score items. 1 = moderate pain; 2 = severe pain; 3 = nausea or vomiting; 4 = feeling dizzy; 5 = shivering; 6 = have been comfortable; 7 = able to mobilise independently; 8 = can hold my baby without assistance; 9 = can feed/nurse baby without assistance; 10 = can look after personal hygiene/toilet; 11 = feeling in control.

*Global health numerical rating scale (NRS), represented as a 100 mm line, marked at each end with anchors 'worst imaginable health state'; to 'best imaginable health state' and with 'sad' or 'happy' stylised representations of faces.

Table 5 Intra-class correlations for ObsQoR-11 items

ObsQoR-11 item	Intra-class correlation
Moderate pain	0.78
Severe pain	0.87
Nausea or vomiting	0.89
Dizziness	0.84
Shivering	0.86
Comfortable	0.62
Mobilisation	0.85
Hold baby	0.72
Nurse baby	0.88
Personal hygiene	0.98
Control	0.85
ObsQoR-11 score	0.96

ObsQoR-11 24 vs 25 h scores. Spearman correlation >0.62 in all items indicates good reliability of the instrument.

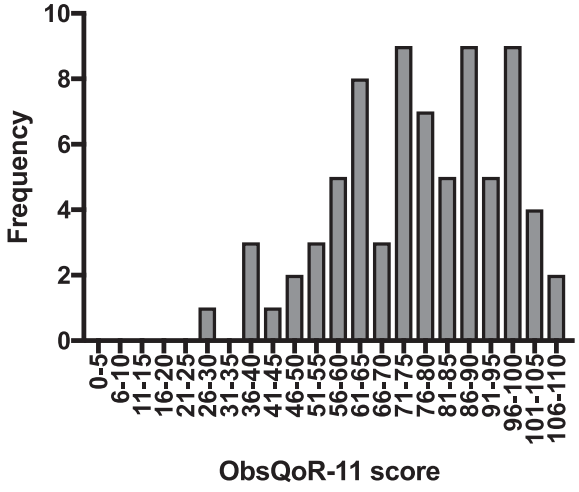


Fig. 2 Histogram of ObsQoR-11 scores 24 h following non-elective caesarean delivery

items, suggesting adequate repeatability and reliability (Table 5). The percentage of women achieving the highest possible ObsQoR-11 score at 24 h was 1% (n=1/100) and for lowest score 0% (n=0/100); therefore, no floor or ceiling effects of the scoring tool were demonstrated. The ObsQoR-11 scores were negatively skewed. The level of skewness was −0.4105 at 24 h postoperatively, indicating that the majority of the ObsQoR-11 scores were in the upper half of the scale (i.e. >55; Fig. 2).

One hundred (100%) women at 24 h and 20 at 25 h following delivery agreed to complete the ObsQoR-11 survey. No patients refused to complete the questionnaire(s). The median [IQR] time taken to complete the ObsQoR-11 questionnaire was 117 [89 to 156] s, with a range from 66 to 300 s.

Discussion

The study evaluated an obstetric-specific recovery tool, the ObsQoR-11, for use in patients undergoing non-elective CD. The ObsQoR-11 performed well in measures of validity, reliability, clinical acceptability and feasibility. Notably, there was a strong correlation of 24-h ObsQoR-11 scores to 24-h global health NRS scores (r = 0.72). This achieved the >0.6 value recommended for health rating scales,[21] which indicates good convergent validity of ObsQoR-11, comparable to the QoR-15 and the more comprehensive QoR-40.[10,14]

The ObsQoR-11 appears to discriminate well between 'good' or 'poor' recovery in patients undergoing non-elective CD, without floor or ceiling effects.[22] A perfect final ObsQoR-11 score is 110, which was reported by only one woman at 24 h postoperatively. A 24-h ObsQoR-11 score ≥96 was associated with a good health status at 24 h (NRS ≥70 mm), and therefore a 'good' recovery, whereas a score of ≤64 was associated with a poorer health status (NRS <70 mm) or 'poor' recovery.[10] In the development study of

ObsQoR-11 involving elective CD, ObsQoR-11 scores also differed significantly for women who had a 'good' versus 'poor' postoperative recovery. The median [IQR] ObsQoR-11 scores in the elective CD population study were 100 [91.3 to 105] versus 87 [72 to 95] ($P < 0.0001$) respectively.[13] The scores in that elective CD population were higher than those of this study, which might be expected given that expedited surgery is associated with factors that could negatively impact recovery such as infection or haemorrhage.[23,24]

Higher ObsQoR-11 scores were associated with a shorter LOS. The median LOS of three nights was greater than our institutional average stay following elective CD (36.5 h or a one- to two-night stay).[13] This was anticipated given the additional adverse factors relating to non-elective CD. Hospital LOS is often used as a measure of in-hospital recovery but may be confounded by local logistics. There may be social reasons why some patients stay longer despite experiencing a 'good' recovery, yet some may self-discharge despite experiencing a 'poor' recovery. Time to readiness for discharge may provide a better index of recovery in future research. The strength of correlation of ObsQoR-11 to LOS was weaker ($r = -0.24$) than that measured for elective CD in our previous study ($r = -0.39$).[13] This could be due to the greater heterogeneity within the non-elective CD population, making prediction more difficult.

The positive correlation of ObsQoR-11 with the parity of women may indicate greater familiarity and less anxiety with the early post-partum period, improving the early recovery experience, although this has not been substantiated.[25] On the other hand, there was no significant correlation among women who had undergone previous CD. We speculate that this may be related to repeat CD being more painful due to scarring, offsetting any theoretical recovery advantage of familiarity.[26] We did not find a significant association with other factors, such as maternal age, BMI, gestational age, gestation, category of CD, duration of surgery, blood loss, or change in haemoglobin. This could be a limitation of the clinical validity of ObsQoR-11 in this setting, or the consequence of few complications or variability in these outcomes in the study cohort, making the study underpowered to detect differences. For example, only 17 women in the study had an estimated blood loss $\geq$1 L and none had a postoperative haemoglobin concentration < 80 g/L. Further use of this scoring tool may clarify or refute these assertions. Internal consistency, as measured using Cronbach's alpha and split-half reliability, achieved the recommended values (0.7–0.9)[21] and values were comparable to those reported for QoR-15[10] and QoR-40.[14] Inter-item correlation also measured internal consistency, with values indicating good correlation of items within the instrument. Reproducibility (measured by test-retest reliability) of the ObsQoR-11 was good, with an $r_i > 0.6$. The short duration (30–60 min) between the test and retest means that any meaningful improvement in recovery is unlikely to have occurred, but test–retest bias and potential selection bias for women perceived as being more willing to repeat the questionnaire cannot be excluded.

We received patient feedback regarding the 'moderate' and 'severe' pain items, which caused some confusion. For example, does any severe pain experienced equate to 10 or 0 on the moderate pain item? It may therefore be advisable to combine these items into a single pain question. Of note, 'moderate pain' had the weakest correlation to global health status NRS and total ObsQoR-11 score of all the items ($r = 0.32$ and 0.35 respectively; Table 2), with 'severe pain' having one of the strongest correlations ($r = 0.51$ and 0.68, respectively). To improve the clinical brevity and clarity of the scoring tool, we therefore suggest that 'moderate pain' be omitted for future studies, which would modify the tool to an amended 'ObsQoR-10'.

We defined postpartum 'functional recovery' as the return to 'baseline' of practical capability after an intervention such as CD. This may be assessed by evaluation of the ability to perform physical acts (such as activities of daily living) at a level deemed appropriate by the patient and physician. Functional recovery may help to evaluate readiness for hospital discharge and the likelihood of re-admission. In obstetrics, functional recovery encompasses both self-care and care of the neonate. Care of the neonate is an important role for new mothers, and aspects of the ObsQoR-11 tool also relate to her ability to self-care (such as comfort, feeling in control and care of personal hygiene). We consider that ObsQoR-11 provides a superior measure of 'functional recovery' compared with other non-obstetric-specific, unidimensional recovery scoring tools.

We acknowledge that the study has several limitations. Responsiveness of ObsQoR-11 was not formally assessed due to the unpredictable nature of non-elective CD, precluding pre-CD ObsQoR-11 score collection. It was measured in the previous study on elective CD by measuring Cohen effect size and change in ObsQoR-11 score from pre- to post-CD, where it was found to have large responsiveness. Serial ObsQoR-11 scores in the early to intermediate postoperative period may provide an alternate measure of change in recovery health status. Recovery scoring tools have been shown to be sensitive in detecting clinically important differences in recovery after elective non-obstetric surgery.[27] Peri-operative interventions that result in a change of 0.9 for the QoR score, 8.0 for QoR-15, or 6.3 for QoR-40, signify a clinically important change which has been demonstrated for non-obstetric surgery and elective CD.[27] An evaluation of the QoR score in detecting a clinically important change following non-elective surgeries is, however, yet to be performed. The 100% response rate at 24 h and the short time taken to com-

plete the questionnaire demonstrate the clinical brevity of the tool. With 20% of women repeating the postoperative test at 25 h, this is comparable to other evaluation studies.[9,10] However, approximately 10% of all women were initially unavailable at the bedside at the 24-h time point of the ObsQoR-11 scoring assessment, but returned to the bedside within one to two hours, which allowed for its completion. The neonatal location (e.g. in the neonatal intensive care unit or with the mother) and neonatal clinical condition was not measured in this study, nor was the duration of labour prior to CD. We acknowledge that these could have a significant impact on maternal reporting of recovery from CD and are therefore limitations of the study. In addition, we did not compare ObsQoR-11 with other methods of quantifying postoperative comfort such as opioid dose or rescue analgesia requirement. Future studies should include these outcomes.

The ObsQoR-11 score has been formulated, evaluated and now validated in a single tertiary referral centre in London, United Kingdom (UK), which may limit the generalisability of our findings nationally and internationally, particularly in institutions with fewer high-risk parturients. We are commencing validation studies at centres in the United States of America and in hospitals serving fewer high-risk pregnancies in the UK. We excluded women who were unable to read or understand written English, so validation is still required in other languages relevant to local demographics.

We measured ObsQoR-11 scores at 24 h as this is becoming a standard recommended by the National Institute for Health and Care Excellence, for discharge from hospital in UK enhanced recovery programmes.[28,29] We did not test ObsQoR-11 beyond 25 h, as has been done for the QoR-15 in non-obstetric day surgery.[9] Future studies should include intermediate and late recovery periods.

In summary, we have evaluated ObsQoR-11 for use in women on the first postoperative day following non-elective CD. The ObsQoR-11 performed well in measures of validity, reliability, clinical acceptability and feasibility. The questionnaire could be used as a standardised, patient-reported outcome measurement tool to evaluate postoperative recovery following CD. Future work is needed to validate ObsQoR-11 after spontaneous and operative vaginal delivery and determine both its global generalisability and its use after 25 h postoperatively. Investigation of the relationship between ObsQoR-11 and other maternal and neonatal outcome measures is also needed.

Disclosure of Interests

None of the authors has any conflicts of interest to declare. This study was internally funded by University College London Hospital.

References

1. Acikel A, Ozturk T, Goker A, et al. Comparison of patient satisfaction between general and spinal anaesthesia in emergency caesarean deliveries. *Turk J Anaesthesiol Reanim* 2017;**45**:41–6.
2. Fatusic Z, Hudic I, Sinanovic O, et al. Short-term postnatal quality of life in women with previous Misgav Ladach caesarean section compared to Pfannenstiel-Dorffler caesarean section method. *J Matern Fetal Neonatal Med* 2011;**24**:1138–42.
3. Deniau B, Bouhadjari N, Faitot V, et al. Evaluation of a continuous improvement programme of enhanced recovery after caesarean delivery under neuraxial anaesthesia. *Anaesth Crit Care Pain Med* 2016;**35**:395–9.
4. Lui MW, Li TKT, Lui F, et al. A randomised, controlled trial of rectus sheath bupivacaine and intrathecal bupivacaine, without or with intrathecal morphine, vs. intrathecal bupivacaine and morphine after caesarean section. *Anaesthesia* 2017;**72**:1225–9.
5. McKeen DM, George RB, Boyd JC, et al. Transversus abdominis plane block does not improve early or late pain outcomes after Cesarean delivery: a randomized controlled trial. *Can J Anaesth* 2014;**61**:631–40.
6. Ma J, Martin R, Chan B, et al. Using activity trackers to quantify postpartum ambulation: a prospective observational study of ambulation after regional anesthesia and analgesia interventions. *Anesthesiology* 2018;**128**:598–608.
7. Gursoy C, Ok G, Aydin D, et al. Effect of anaesthesia methods for regaining daily life activities in cesarean patients. *Turk J Anaesthesiol Reanim* 2014;**42**:71–9.
8. Trivino-Juarez JM, Romero-Ayuso D, Nieto-Pereda B, et al. Health related quality of life of women at the sixth week and sixth month postpartum by mode of birth. *Women Birth* 2017;**30**:29–39.
9. Chazapis M, Walker EM, Rooms MA, et al. Measuring quality of recovery-15 after day case surgery. *Br J Anaesth* 2016;**116**:241–8.
10. 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.
11. Jansen AJ, Essink-Bot ML, Duvekot JJ, et al. Psychometric evaluation of health-related quality of life measures in women after different types of delivery. *J Psychosom Res* 2007;**63**:275–81.
12. Petrou S, Kim SW, McParland P, et al. Mode of delivery and long-term health-related quality-of-life outcomes: A prospective population-based study. *Birth* 2017;**44**:110–9.
13. Ciechanowicz S, Setty T, Robson E, et al. Development and evaluation of an obstetric quality-of-recovery score (ObsQoR-11) after elective Caesarean delivery. *Br J Anaesth* 2019;**122**:69–78.
14. Myles PS, Weitkamp B, Jones K, et al. Validity and reliability of a postoperative quality of recovery score: the QoR-40. *Br J Anaesth* 2000;**84**:11–5.
15. Royal College of Obstetricians and Gynaecologists and The Royal College of Anaesthetists: Classification of Urgency of Caesarean Section – A Continuum of Risk. April 2010. Available at: https://www.rcog.org.uk/globalassets/documents/guidelines/goodpractice11classificationofurgency.pdf. Accessed 27 December, 2018.
16. Wei J, Wu QJ, Zhang TN, et al. Complications in multiple gestation pregnancy: a cross-sectional study of ten maternal-fetal medicine centers in China. *Oncotarget* 2016;**7**:30797–803.
17. DePasse JM, Durand W, Daniels AH. Predictors of non-neurologic complications and increased length of stay after cervical spine osteotomy. *World Neurosurg* 2018;**118**:e727–30.
18. Musallam KM, Tamim HM, Richards T, et al. Preoperative anaemia and postoperative outcomes in non-cardiac surgery: a retrospective cohort study. *Lancet* 2011;**378**:1396–407.
19. Horie S, Nomura K, Takenoshita S, et al. A relationship between a level of hemoglobin after delivery and exclusive breastfeeding initiation at a baby friendly hospital in Japan. *Environ Health Prev Med* 2017;**22**:40.

20. Lj C. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16**:297–334.
21. McDowell INC. Measuring health: a guide to rating scales and questionnaires. New York: Oxford University Press; 2006.
22. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;**60**:34–42.
23. Wang J, Ma R, Churilov L, et al. The cost of perioperative complications following pancreaticoduodenectomy: a systematic review. *Pancreatology* 2018;**18**:208–20.
24. Patel AS, Bergman A, Moore BW, et al. The economic burden of complications occurring in major surgical procedures: a systematic review. *Appl Health Econ Health Policy* 2013;**11**:577–92.
25. Pereira TRC, Souza FG, Beleza ACS. Implications of pain in functional activities in immediate postpartum period according to the mode of delivery and parity: an observational study. *Braz J Phys Ther* 2017;**21**:37–43.
26. Beucher G, Dolley P, Levy-Thissier S, et al. Maternal benefits and risks of trial of labor versus elective repeat caesarean delivery in women with a previous caesarean delivery. *J Gynecol Obstet Biol Reprod (Paris)* 2012;**41**:708–26.
27. Myles PS, Myles DB, Galagher W, et al. Minimal clinically important difference for three Quality of Recovery scales. *Anesthesiology* 2016;**125**:39–45.
28. Wrench IJ, Allison A, Galimberti A, et al. Introduction of enhanced recovery for elective caesarean section enabling next day discharge: a tertiary centre experience. *Int J Obstet Anesth* 2015;**24**:124–30.
29. NICE. Caesarean Section 2011. Available at: https://www.nice.org.uk/guidance/cg132. Accessed 27 December, 2018.