BMJ Open

Psychometric validation of the Coronary Revascularisation Outcome Questionnaire (CROQv2) in the context of the NHS Coronary Revascularisation PROMs Pilot

Sara Schroter,[1,2] Rebecca Miles,[3] Stephen Green,[3] Mark Jackson[4]

To cite: Schroter S, Miles R, Green S, *et al*. Psychometric validation of the Coronary Revascularisation Outcome Questionnaire (CROQv2) in the context of the NHS Coronary Revascularisation PROMs Pilot. *BMJ Open* 2017;**7**:e015915. doi:10.1136/bmjopen-2017-015915

► Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2017-015915).

Received 9 January 2017
Accepted 2 February 2017

For numbered affiliations see end of article.

Correspondence to
Dr Sara Schroter;
sschroter@bmj.com

ABSTRACT

Objectives: The Coronary Revascularisation Outcome Questionnaire (CROQ) is a patient-reported outcome measure (PROM) for coronary artery bypass surgery (CABG) and percutaneous coronary intervention (PCI). We tested the psychometric properties of a modified version (CROQv2) when administered in a National Health Service (NHS)/Department of Health (DH) funded pilot of PROMs for coronary revascularisation.

Design: Psychometric validation study.

Setting: 11 English hospitals in the UK taking part in the NHS/DH funded pilot of PROMs for coronary revascularisation.

Participants: Comprehensive analyses of acceptability, reliability, validity and responsiveness were conducted independently for each of the prerevascularisation (n=2685 and n=3711) and postrevascularisation (n=869 and n=837) versions of the CROQ-CABG and CROQ-PCI, respectively.

Results: All versions met prespecified stringent criteria for (1) acceptability of items (missing data) and scales (missing data, floor and ceiling effects, skewness); (2) tests of scaling assumptions; (3) reliability: internal consistency (Cronbach's α, item-total correlations); (4) construct validity based on within-scale analyses (internal consistency, intercorrelations between scales, factor analysis and hypothesis testing); (5) construct validity based on comparisons with external measures (convergent and discriminant validity and hypothesis testing) and (6) responsiveness. Results were also confirmed when tests were repeated on subsamples of CABG (n=639) and PCI (n=615) patients who reported receiving help completing prerevascularisation questionnaires.

Conclusions: The availability of a psychometrically robust procedure-specific tool that could be used as part of a large-scale coronary revascularisation PROMs programme to capture the patients' perspective of coronary revascularisation will enable outcomes important to patients to be routinely collected alongside clinical outcomes. The CROQ is suitable for administration by postal survey or the prerevascularisation versions can be administered in the clinical setting as in the Coronary Revascularisation PROMs Pilot.

Strengths and limitations of this study

- The Coronary Revascularisation Outcome Questionnaire includes a much broader range of outcomes important to patients than other cardiac-specific questionnaires.
- The availability of a tool suitable for use in large-scale patient-reported outcome measures (PROMs) programmes, alongside the collection of clinical data, could enable the routine collection of outcomes that matter to coronary revascularisation patients, rather than focus on narrow aspects of disease or functioning.
- Psychometric validation is an iterative process; it is from repeated use in large samples that we can gain confidence that PROMs are measuring what they intend to measure in a reliable way and are able to detect important change.
- Large-scale psychometric validation of a procedure-specific PROM for coronary revascularisation in 11 hospitals in the UK.
- Unable to measure test–retest reliability.

INTRODUCTION

Patient-reported outcome measures (PROMs) measure health status and health-related quality of life (HRQoL) from the patients' perspective. There is growing interest in capturing PROMs data for patients with coronary heart disease and other cardiac conditions and there are numerous disease-specific tools available.[1] However, most have been developed to evaluate HRQoL in medically rather than surgically treated patients, and many have not been rigorously validated.[2] The Coronary Revascularisation Outcome

Questionnaire (CROQ)[3 4] is a PROM to evaluate health status and HRQoL in patients undergoing coronary artery bypass graft surgery (CABG) and percutaneous coronary intervention (PCI). It was developed and validated in 2000, as a self-administered survey tool to evaluate outcomes in research and clinical audit, at prerevascularisation and 3-months postrevascularisation. It is currently the only disease-specific tool developed specifically to measure health outcomes before and after coronary revascularisation with some demonstrated evidence of reliability, validity and responsiveness, but has not been used widely.

PROMs are most commonly used to measure the impact of healthcare in research and audit but in recent years have been used to compare the performance of healthcare providers.[5] Since 2009, the National Health Service (NHS) in England has used PROMs to assess outcomes in four elective surgical procedures (hip and knee replacement, varicose vein surgery and groin hernia surgery) on a routine basis for the purpose of service evaluation.[5–7] NHS England and the Department of Health (DH) also use the data to monitor progress towards strategic objectives, such as those specified in the NHS Outcomes Framework.[8] The NHS Coronary Revascularisation PROMs Pilot was launched in November 2011 in order to evaluate the feasibility of extending the NHS PROMs programme to patients before and after coronary revascularisation.[9] The pilot was established to examine the feasibility of collecting PROMs data from patients selected for elective first-time coronary revascularisation across 11 hospitals in England. Consistent with the PROMs being collected routinely in other procedures[5–7] the DH chose to use the generic EQ-5D-3L[10] alongside a procedure-specific instrument. The CROQ[3 4] was chosen as the procedure-specific instrument by the DH despite it not having been used widely, as it had demonstrated preliminary evidence of psychometric robustness, included a broader range of outcomes than other cardiovascular-specific measures, and was specific to coronary revascularisation.[2]

It is essential that PROMs satisfy certain development, psychometric and scaling standards if they are to provide reliable and valid information for decision making. Classical Test Theory (CTT),[11–15] the traditional psychometric approach, is the most dominant paradigm in the development of PROMs.[16] Within CCT, well-established methods and criteria are applied to indicate that the concept, that is, represented by the PROM is clear and well understood; that the content is relevant to the patient group concerned; that the psychometric properties (acceptability, reliability, validity, responsiveness) are adequate; and that the scaling structure is justified.[17–19] The psychometric properties of a PROM are sample and context dependent within the CTT paradigm.[18 19] Psychometric validation is an iterative process;[17] it is from repeated use in large samples that we can gain confidence that PROMs are measuring what they intend to measure in a reliable way and are able to detect important change.[18 19]

Psychometric properties are influenced by how people respond and this can be influenced by many factors including use in different patient groups, a change in the mode of administration or setting, a change in the assessment point, minor changes to phrasing, question order or response format.[18–23] Completion of PROMs by persons other than the patient can introduce bias; for example, family and professionals tend to underestimate patient's quality of life status across diverse cultures and health conditions.[24] All changes made to the original version of an instrument require revalidation of the instrument, when used in a new context.[18] For use in the NHS Coronary Revascularisation PROMs Pilot, changes were made to the original version of the CROQ (necessitating a new version, CROQv2), including some changes to the text, a change to the postrevascularisation assessment point (from 3 to 6 months postrevascularisation), method of administration, setting of administration and sampling frame (see table 1 and online supplementary appendices 1–4). The NHS Coronary Revascularisation PROMs Pilot, also included a much larger, and potentially wider group of patients in terms of demographic profile and case mix to the sample used to validate CROQv1. These changes necessitated a re-evaluation of the psychometric properties of the CROQ in the new context. We describe the psychometric validation of CROQv2 in the context of the NHS Coronary Revascularisation PROMs Pilot.

METHODS

Samples

The samples described in this paper are those used for the psychometric validation only and are a subset of all the patient data gathered in the main NHS Coronary Revascularisation PROMs Pilot.

Prerevascularisation (Q1) samples

All NHS PROMs Pilot patients waiting for coronary revascularisation who completed a prerevascularisation (Q1) questionnaire were included except those categorised as ineligible (n=63) or duplicates (n=31). A total of 6396 (2685 CABG and 3711 PCI) patients were included.

Postrevascularisation (Q2) samples

Patients were only sent a postrevascularisation (Q2) questionnaire if they had already completed a Q1 questionnaire. For consistency with the rest of the NHS PROMs programme,[5 6] the DH intended that the postrevascularisation Q2 questionnaire would be sent to patients at 6 months post procedure in the NHS Coronary Revascularisation PROMs Pilot. However, in practice the Q2 was not administered at a fixed interval of 6 months after the revascularisation date, and there was wide variation in the interval between the revascularisation procedure and Q2 completion. For the Q2 psychometric analysis, only patients who completed a Q2

Table 1 Changes made to the CROQ instrument or administration of it

Type of change	CROQv1[3]	NHS Coronary Revascularisation PROMs Pilot
Textual changes	–	Owing to changes in clinical practice over time, PCI is now often performed via the radial artery rather than the femoral artery in the groin. Textual changes were necessary for all items in the CROQ-PCI adverse effects scale.
Assessment point	Validated for use with patients at pre-coronary and 3-months post-coronary revascularisation.	The postrevascularisation assessment point (Q2) was changed to 6 months postrevascularisation.
Method of administration	Prerevascularisation and postrevascularisation versions were administered by postal survey.	Prerevascularisation questionnaires were administered in the clinical setting by staff with assistance provided as required. The reason for requiring assistance will have varied, for example, some patients might not have been able to read English and others might just have needed help interpreting a single item.
Setting	Prerevascularisation and postrevascularisation versions administered by postal survey to patients' home addresses. This enabled patients to complete the questionnaires at leisure and privately outside of the clinical setting.	Prerevascularisation questionnaires were administered in the clinical setting with assistance provided by staff for those requiring it.
Sampling frame	English speaking patients undergoing elective coronary revascularisation procedures at three hospitals in the UK.	Patients undergoing elective procedures at 11 hospitals in the UK with a potentially different case mix severity and demographic profile. Patients speaking all languages were offered the prerevascularisation questionnaire by staff in the clinical setting. Staff offered to assist patients, for example, by reading the questions to patients and explaining the meaning and a telephone translation service was available on request.

CROQ, Coronary Revascularisation Outcome Questionnaire; PCI, percutaneous coronary intervention.

questionnaire within 5–7 months of a linked Hospital Episode Statistics (HES) revascularisation episode were included. The Q2 psychometric samples therefore included 869 CABG and 837 PCI patients.

Responsiveness samples

Slightly more stringent inclusion criteria were applied to the responsiveness psychometric samples to evaluate whether the CROQ was sensitive to change (between Q1 and Q2) in the context of a specific single elective procedure. Patients who had an emergency or a repeat elective coronary revascularisation procedure were excluded. The responsiveness samples included 865 CABG and 811 PCI patients.

Psychometric evaluation

There is a prerevascularisation and postrevascularisation version for CABG and PCI. Each of the four versions are scored to produce four core scales: symptoms, physical functioning, psychosocial functioning and cognitive functioning. In addition, the postrevascularisation versions include additional items that are scored to produce the satisfaction and adverse effects scales. All scales are scored on a 0–100 scale with higher scores reflecting better functioning.[3]

We evaluated the acceptability, reliability, validity and responsiveness of CROQv2 prerevascularisation and postrevascularisation versions independently in CABG and PCI samples using the same widely adopted criteria, based on CTT, used to validate CROQv1.[3] Table 2 provides an overview of the tests and criteria applied. Analyses included an evaluation of (1) the acceptability of items (missing data) and scales (missing data, floor and ceiling effects, skewness); (2) tests of scaling assumptions; (3) reliability: internal consistency (Cronbach'sα, item-total correlations); (4) construct validity based on within-scale analyses (internal consistency, intercorrelations between scales, factor analysis and hypothesis testing); (5) construct validity based on comparisons with external measures (convergent and discriminant validity and hypothesis testing) and (6) responsiveness.

In addition, the same prerevascularisation psychometric tests were applied to the subgroups of patients (n=639 CABG and n=615 PCI) who reported they received help completing the Q1 questionnaire in the clinical setting to see if the psychometric properties of the CROQ were compromised. The provision of help with questionnaire completion to respondents might have enabled a wider group of patients to have been included in this sample, than in the self-completed only sample. The type of help received and from whom was not recorded.

Table 2 Psychometric tests and criteria*

Psychometric property	Definition/test	Criteria
Acceptability	Quality of data; assessed by completeness of data and score distributions.	▶ Proportion of missing data for scales (<10%)[3] ▶ Low floor/ceiling effects in the pre-revascularisation samples (percentage scoring lowest/highest possible scale scores)
Reliability: internal consistency	Extent to which items in a scale measure the same construct (such as homogeneity of the scale); assessed by Cronbach's α, item-total correlations, and value of α if an item is deleted from a scale.	▶ Cronbach's α for scales >0.70[13] ▶ Item-total correlations >0.30[13] ▶ Value of α if an item is deleted from scale should not 'substantially increase'[25]
Tests of scaling assumptions	Evidence that an item belongs in its own scale and not another scale (item convergent and discriminant validity).	▶ Scaling success/failure (item does/does not correlate significantly higher with own scale than other scale) and probable scaling success/failure (item does/does not correlate more highly, but not significantly, with own scale than other scales)[26]
Construct validity (within scale analyses)	Evidence that each scale measures a single construct and that items can be combined to form scales; assessed on the basis of evidence of good internal consistency, factor analysis and correlations between scale scores.	▶ Internal consistency (Cronbach's α >0.70) ▶ Principle axis factor analysis (factor loadings ≥0.30)[3] ▶ Moderate intercorrelations between scale scores and evidence of unique reliable variance (reliability coefficients with values greater than the intercorrelations between scales)[26]
Construct validity (analyses against external criteria): convergent and discriminant validity	Evidence that scales are correlated with other measures of the same or similar construct, and not correlated with other measures of different constructs; assessed on the basis of correlations between CROQ, EQ-5D-3L, and age and sex.	▶ Magnitude and direction of correlations expected to vary according to the similarity of constructs being measured in each instrument ▶ Low to moderate correlations expected between a disease-specific and generic tool ▶ Very low correlations (<0.30) expected for age and sex
Construct validity (analyses against external criteria): hypothesis testing	Evidence that scales differentiate known groups; assessed by comparing CROQ scores between groups hypothesised to differ.	▶ CROQ scores should be significantly (p<0.05) different for groups expected to differ
Responsiveness	Ability of scales to detect clinically important change over time between Q1 and Q2. Assessed by effect sizes (mean change score between prerevascularisation and postrevascularisation divided by the SD of scores at pre-revascularisation).	▶ Effect sizes defined as small (0.20), medium (0.50) and large (≥0.80)[27]

*Adapted from Schroter and Lamping 2004.[3]
CROQ, Coronary Revascularisation Outcome Questionnaire.

RESULTS

Table 3 shows the respondent characteristics for the main psychometric samples. There was a higher proportion of male patients (86.0% vs 83.6%) and white patients (79.6% vs 59.9%) in the CABG Q2 sample than in the CABG Q1 sample. There was a higher proportion of male patients (77.9% vs 74.9%) and white patients (85.9% vs 61.1%) in the PCI Q2 sample than the PCI Q1 sample. PCI patients completing Q2 were also older on average than those completing Q1 (mean=66.0 (SD 9.7)) versus 64.7 (SD 10.7) years. A higher proportion of patients in the CABG (23.8%) and PCI (16.6%) Q1 samples reported that they had help completing their questionnaires than patients in the CABG (11.4%) and PCI (8.6%) Q2 samples. This probably reflects the fact that hospital staff was at hand to help patients complete questionnaires in the clinical setting at pre-revascularisation.

Table 3 Respondent characteristics: CABG and PCI samples

	CABG prerevascularisation (Q1) sample (N=2685)*	CABG postrevascularisation (Q2) sample (N=869)†	PCI prerevascularisation (Q1) sample (N=3711)*	PCI postrevascularisation (Q2) sample (N=837)†
Mean (SD) age in years	66.4 (9.7)	66.72 (8.7)	64.7 (10.7)	66.0 (9.7)
Median (LQ, UQ) age in years	67.0 (60.0, 74.0)	67.0 (61.0, 73.0)	65.0 (57.0, 73.0)	66.0 (59.0, 74.0)
Number (proportion) males	2245 (83.6)	747 (86.0)	2778 (74.9)	652 (77.9)
Number (proportion) living with partner/spouse/family/friends	2161 (80.5)	706 (81.2)	2970 (80.0)	695 (83.0)
Number (proportion) living alone	468 (17.4)	146 (16.8)	641 (17.3)	130 (15.5)
Number (proportion) considering themselves to have a disability	785 (29.2)	181 (20.8)	1340 (36.1)	261 (31.2)
Number (proportion) receiving help completing questionnaire	639 (23.8)	99 (11.4)	615 (16.6)	72 (8.6)
Number (proportion) of each ethnic group‡				
White	1608 (59.9)	692 (79.6)	2268 (61.1)	719 (85.9)
Mixed	9 (0.3)	4 (0.5)	12 (0.3)	6 (0.7)
Asian	178 (6.6)	64 (7.4)	156 (4.2)	24 (2.9)
Black	9 (0.3)	0 (0)	15 (0.4)	2 (0.2)
Other ethnic groups	34 (1.3)	13 (1.5)	33 (0.9)	12 (1.4)
Missing	847 (31.5)	96 (11.0)	1227 (33.1)	74 (8.8)

Numbers in brackets are per cents, unless specified otherwise. Numbers do not sum to 100% due to rounding or missing data.
*All patients who completed Q1 regardless of whether they had a HES linked episode or whether they completed Q2.
†Patients who completed only one Q1 questionnaire, had a matched HES linked episode based on original PROMS programme criteria, were sent a Q2 at a time based on the HES linked episode, and completed Q2 within 5–7 months of the HES linked episode.
‡http://www.datadictionary.nhs.uk/data_dictionary/attributes/e/end/ethnic_category_code_de.asp?shownav=1
CABG, coronary artery bypass surgery; LQ, lower quartile; PCI, percutaneous coronary intervention; UQ, upper quartile.

Acceptability

There was a low level of missing data across all items in each version (prerevascularisation and postrevascularisation); all scales at prerevascularisation and postrevascularisation had <3% missing data (table 4). Analysis of missing item data in the more elderly subsamples did not suggest that it was overly burdensome for elderly patients. Scale scores were calculated with a possible score range of 0–100, as described for CROQv1.[3] There were no floor effects (a high proportion scoring at the bottom of the scales) in the prerevascularisation or postrevascularisation samples, but there were ceiling effects (a high proportion scoring at the top of the scales) in the postrevascularisation samples, as expected following effective interventions and small ceiling effects for Cognitive Functioning at prerevascularisation.

Reliability

Cronbach's α coefficients for all scales at prerevascularisation and postrevascularisation far exceeded the criterion of >0.70[13] indicating excellent internal consistency (table 4). For all scales in all samples, the value of α if item was deleted did not substantially increase indicating that all of the items within each scale were contributing to the underlying constructs.[25] Scales in all versions demonstrated evidence of homogeneity. All item-total correlations exceeded the criterion of >0.30,[13] the range of item-total correlations was small to moderate, and the average item-total correlations were moderate to high.

Tests of scaling assumptions

The results of these tests provided strong confirmatory evidence that the CROQ items are correctly grouped in scales in the prerevascularisation and postrevascularisation versions (table 4).[26]

Construct validity (within scale analyses)

Construct validity was demonstrated by evidence of high internal consistency (high values of Cronbach's α and moderately high item-total correlations, table 4). Principal axis factor analysis and the pattern of intercorrelations between the CROQ scales confirmed the scaling structure of each version (data not presented). Analyses of CROQ scale scores showed the expected pattern for groups hypothesised to differ. For example, CABG and PCI patients who reported overall improvement in their heart condition at Q2 scored significantly higher ($p<0.05$) on all four CROQ scales than those who reported their heart condition as being the same or worse at Q2 (data not presented).

Construct validity (analysis against external criteria)

Convergent and discriminant validity

CROQ scales at prerevascularisation and postrevascularisation were more highly correlated with the EQ-5D-3L[10] dimensions measuring conceptually similar constructs than with those measuring different constructs (table 5). The correlation coefficients were low to moderate as

Table 4 Acceptability, reliability and tests of scaling assumptions for CROQ-CABG and CROQ-PCI

CROQ scale	Range of scores		Acceptability		Internal consistency			Tests of scaling assumptions
	Scale	Sample	% missing	% floor/% ceiling	n	Cronbach's α	Mean item-total correlation	% definite scaling success (% probable scaling success)
CABG prerevascularisation (N=2685)								
Symptoms	0–100	0–100	2.3	0.3/5.5	2622	0.87	0.65	100 (0)
Physical functioning	0–100	0–100	2.1	2.0/10.9	2628	0.91	0.71	100 (0)
Psychosocial functioning	0–100	0–100	1.2	0.1/1.5	2653	0.94	0.72	100 (0)
Cognitive functioning	0–100	0–100	1.2	1.5/30.0	2654	0.90	0.79	100 (0)
PCI prerevascularisation (N=3711)								
Symptoms	0–100	0–100	2.6	0.3/4.2	3616	0.89	0.69	100 (0)
Physical functioning	0–100	0–100	2.1	3.7/10.0	3632	0.92	0.74	100 (0)
Psychosocial functioning	0–100	0–100	1.2	0.5/1.7	3665	0.95	0.74	100 (0)
Cognitive functioning	0–100	0–100	1.2	1.8/28.4	3667	0.91	0.81	100 (0)
CABG postrevascularisation (N=869)								
Symptoms	0–100	6–100	2.2	0.0/33.1	850	0.88	0.67	100 (0)
Physical functioning	0–100	0–100	2.0	0.7/39.8	852	0.93	0.76	100 (0)
Psychosocial functioning	0–100	0–100	0.5	0.1/13.2	865	0.96	0.77	100 (0)
Cognitive functioning	0–100	0–100	0.6	0.3/41.4	864	0.91	0.82	100 (0)
Adverse effects	0–100	0–100	0.6	0.1/13.1	864	0.86	0.56	100 (0)
Satisfaction	0–100	0–100	0.1	0.1/33.3	868	0.79	0.57	100 (0)
PCI postrevascularisation (N=837)								
Symptoms	0–100	0–100	2.6	0.1/23.4	815	0.91	0.73	100 (0)
Physical functioning	0–100	0–100	2.4	1.4/35.5	817	0.94	0.80	100 (0)
Psychosocial functioning	0–100	5–100	0.5	0.0/11.0	833	0.96	0.78	93 (7)
Cognitive functioning	0–100	0–100	0.8	0.6/36.9	830	0.90	0.80	100 (0)
Adverse effects	0–100	0–100	1.3	0.1/70.3	826	0.84	0.62	100 (0)
Satisfaction	0–100	0–100	0.2	0.5/26.9	835	0.83	0.61	100 (0)

CABG, coronary artery bypass surgery; CROQ, Coronary Revascularisation Outcome Questionnaire; PCI, percutaneous coronary intervention.

Table 5 Construct validity (convergent and discriminant validity): CROQ-CABG and CROQ-PCI correlations with EQ-5D-3L

CROQ scale	EQ-5D-3L dimensions					EQ-5D-3L VAS score
	Mobility	Self-care	Usual activities	Pain and discomfort	Anxiety and depression	
CABG prerevascularisation (n=2685)						
Symptoms	−0.41	−0.26	−0.42	**−0.59**	−0.27	0.47
Physical functioning	**−0.58**	**−0.40**	**−0.55**	−0.49	−0.27	0.52
Psychosocial functioning	−0.40	−0.29	−0.52	−0.46	**−0.57**	0.55
Cognitive functioning	−0.32	−0.29	−0.38	−0.35	**−0.44**	0.41
PCI prerevascularisation (n=3711)						
Symptoms	−0.46	−0.30	−0.46	**−0.59**	−0.32	0.50
Physical functioning	**−0.62**	**−0.47**	**−0.57**	−0.50	−0.31	0.56
Psychosocial functioning	−0.48	−0.42	−0.57	−0.48	**−0.57**	0.59
Cognitive functioning	−0.36	−0.40	−0.41	−0.38	**−0.49**	0.45
CABG postrevascularisation (N=869)						
Symptoms	−0.43	−0.37	−0.52	**−0.50**	−0.47	0.54
Physical functioning	**−0.63**	**−0.55**	**−0.68**	−0.52	−0.45	0.60
Psychosocial functioning	−0.50	−0.51	−0.64	−0.56	**−0.68**	0.65
Cognitive functioning	−0.41	−0.47	−0.53	−0.46	**−0.54**	0.50
Adverse effects	−0.35	−0.40	−0.45	−0.55	−0.44	0.45
Satisfaction	−0.34	−0.26	−0.48	−0.44	−0.44	0.52
PCI postrevascularisation (N=837)						
Symptoms	−0.53	−0.38	−0.56	**−0.61**	−0.44	0.60
Physical functioning	**−0.71**	**−0.57**	**−0.71**	−0.50	−0.41	0.68
Psychosocial functioning	−0.60	−0.51	−0.66	−0.56	**−0.65**	0.67
Cognitive functioning	−0.49	−0.41	−0.52	−0.40	**−0.56**	0.53
Adverse effects	−0.25	−0.25	−0.23	−0.26	−0.18	0.21
Satisfaction	−0.35	−0.27	−0.42	−0.41	−0.34	0.47

Values in bold indicate correlations between scores that purport to measure similar aspects of HRQoL.
The correlations with the EQ-5D-3L dimension scores are negative as higher scores on these indicate poor functioning whereas higher scores on the CROQ and EQ-5D-3L VAS score indicate better functioning.
CABG, coronary artery bypass surgery; CROQ, Coronary Revascularisation Outcome Questionnaire; HRQoL, health-related quality of life; PCI, percutaneous coronary intervention.

expected between a disease-specific and generic tool, and slightly higher at postrevascularisation. As hypothesised there were very low correlations (all <0.30) between each version of the CROQ and age and sex, demonstrating that scores are not biased by these demographic factors.

Hypothesis testing/known groups (analyses against external criteria)

The EQ-5D-3L User Guide advises that levels 2 (some problems) and 3 (extreme problems) are combined into a single category (problems) as extreme problems are often low in frequency. While extreme problems were reported by some patients for some EQ-5D dimensions, the numbers were small for this level for other dimensions so the levels were collapsed for all dimensions. As hypothesised, mean CROQ scale scores were significantly higher (p<0.001) for those reporting 'no problems' than 'problems' on the five EQ-5D-3L dimensions, at prerevascularisation and postrevascularisation in the CABG and PCI samples (web tables S1 and S2). In addition, on average, CABG and PCI patients reporting a comorbidity of depression scored significantly lower (p<0.05) on all CROQ scales, at prerevascularisation and postrevascularisation.

Responsiveness

Table 6 shows the effect sizes for change between prerevascularisation and postrevascularisation for the four core scales in the CABG and PCI responsiveness samples. All scales demonstrated significant change between prerevascularisation and postrevascularisation (p<0.001). For the CROQ-CABG, there was a large[27] effect size for symptoms and psychosocial functioning, a moderate effect size for physical functioning, and a small effect size for cognitive functioning. For the CROQ-PCI, there was a large effect size for symptoms, moderate effect sizes for physical functioning and psychosocial functioning and a very small effect size for cognitive functioning. In the CABG and PCI samples, the generic EQ-5D-3L visual analogue scale (VAS) score had a smaller effect size (was less responsive) than three of the four disease-specific CROQ scales.

Subsamples of patients reporting having received help with the Q1 questionnaire

A significantly higher proportion of patients receiving help completing the Q1 questionnaire compared with those who did not receive help were female (CABG: 21.3% vs 14.5%; PCI: 30.4% vs 23.9%, p=.001) and

Table 6 Responsiveness of CROQv2 from prerevascularisation to postrevascularisation

CROQ scale	Responsiveness effect size
CABG responsiveness sample (n=865)	
CROQ symptoms	1.06
CROQ physical functioning	0.72
CROQ psychosocial functioning	0.91
CROQ cognitive functioning	0.28
EQ-5D-3L VAS score	0.71
PCI responsiveness sample (n=811)	
CROQ symptoms	0.82
CROQ physical functioning	0.52
CROQ psychosocial functioning	0.59
CROQ cognitive functioning	0.19
EQ-5D-3L VAS score	0.42

CABG, coronary artery bypass surgery; CROQ, Coronary Revascularisation Outcome Questionnaire; PCI, percutaneous coronary intervention.

considered themselves to have a disability (CABG 41.8% vs 26.4%; PCI: 58.4% vs 33.3%, p<0.001). A significantly lower proportion of CABG patients receiving help were white (49.8% vs 63.1%, p<0.001). CABG patients (68.6 (SD 9.9) years vs 65.6 (SD 9.5) years) and PCI patients (68.4 (SD 10.9) years vs 63.85 (SD 10.5) years) who received help were also significantly older (p<0.001) and scored significantly lower on all four core scales of the CROQ and the EQ-5D-3L VAS Score (p<0.001), than those who did not receive help. All tests of acceptability, scaling assumptions, reliability and validity met the same psychometric criteria when they were repeated for just the subsamples of patients who reported that they received help completing the prerevascularisation versions in the clinical setting (n=639 CABG and n=615 PCI).

DISCUSSION

Traditional psychometric properties of PROMs are context and sample dependent.[18 19] Traditional psychometric analyses showed that the prerevascularisation and postrevascularisation versions of the CROQ-CABG and CROQ-PCI demonstrated sufficient evidence of acceptability, reliability (internal consistency), validity and responsiveness when used in the context of the NHS Coronary Revascularisation PROMs Pilot, for the sample of patients whose postprocedure assessment point was fixed to between 5 and 7 months of a HES confirmed revascularisation procedure. Analyses also confirmed that the prerevascularisation versions of CROQv2 are robust for self-completion and for completion with patient-reported help when administered in a clinical setting.

The initial psychometric validation of CROQv1 showed it to be acceptable, reliable, valid and responsive when administered via postal survey, in the context of a research project, in selected samples of patients at prerevascularisation and 3 months postrevascularisation.[3 4] The analysis described in this paper confirms that the psychometric properties for CROQv2 are also robust when it is administered to a more diverse and larger number of patients undergoing elective coronary revascularisation procedures, outside of the context of a research project. The CROQ was developed as a self-administered postal survey, but our subgroup analysis demonstrated that the psychometric properties were not compromised when patients received help in the clinical setting, despite the fact that research has shown that family and professionals tend to underestimate patients quality of life status across diverse cultured and health conditions.[24] Our analysis also confirmed that the psychometric properties are withheld when the postrevascularisation version is administered at 5–7 months (rather than 3 months) postrevascularisation by postal survey. As such it is appropriate to administer the CROQ at prerevascularisation (by survey or in the clinical setting) and at 3 or 6 months postrevascularisation by postal survey. This will allow for a greater degree of flexibility in future study designs and may reduce administration costs.

The CROQ is the only validated disease-specific PROM developed specifically to measure outcomes before and after coronary revascularisation (CABG and PCI). While other cardiac-specific PROMs have been developed and are relevant for use with coronary heart disease patients, such as the Seattle Angina Questionnaire,[28] the MacNew Heart Disease Health-Related Quality of Life Questionnaire,[29] and the Quality of Life Index-Cardiac Version, QLI-CV,[30] these questionnaires do not capture all outcomes of importance to patients before and after coronary revascularisation.[2] While some of these questionnaires have been widely used with coronary heart disease patients, including those undergoing CABG and PCI, for example, the Seattle Angina Questionnaire, when selecting an instrument it is important to ensure that the most relevant and applicable PROM is used for the research question under study, that all the questions are applicable to the specific patient group and that items of importance to patients are included.

Study limitations

This study has some important limitations. First, a large number of patients had to be excluded from the main NHS PROMs Pilot postrevascularisation samples, as the interval when patients were sent their postoperative questionnaires (Q2) for completion at home, was very varied. As PROMs are sample and context dependent, to perform meaningful psychometric analysis, it was essential to compare patients at a similar point in time after revascularisation. The possibly slightly lenient criterion of including patients who completed their postoperative Q2 questionnaire between 5 and 7 months of a HES confirmed coronary revascularisation date was applied to all the postrevascularisation psychometric analysis. In future applications, if these essential exclusion criteria are not applied then the psychometric properties of the

CROQ may be compromised and the data may be invalid.

Second, it was not possible to evaluate the stability of the CROQv2 through test–retest reliability[17] as the appropriate data was not collected during the NHS PROMs Revascularisation Pilot. This should be assessed in a small random sample of CABG and PCI patients, if the decision to use CROQv2 more widely in this context is made.

Third, at the time the CROQv1 was originally developed, the dominant psychometric paradigm was CTT and the CROQv1 was developed using these traditional methods (as described here). It was therefore important to assess the psychometric properties of CROQv2 using the same methods as the original validation. However, future work could evaluate the CROQv2 using so called modern psychometric methods such as Item Response Theory[31] or Rasch Measurement Theory.[32] This would enable CROQv2 scores to be placed on a truly interval scale, to be invariant (ie, independent of sample and context) and potentially to be applicable in clinical practice at the individual patient level.[33] Currently, CROQv2 should not be used as a tool to assess a patient's need for surgery as, like other PROMs, it has not been validated for this purpose and it is possible that its predictive validity is not strong enough.

Conclusion

The CROQ is reliable, valid and responsive when used in the context of a large-scale PROMs programme. While there are several validated cardiac specific PROMs, the CROQ remains the only validated procedure-specific questionnaire for coronary revascularisation. It was developed with patients and includes a much broader range of outcomes important to patients than other cardiac-specific questionnaires. The availability of a tool suitable for use in large-scale PROMs programmes, alongside the collection of clinical data, could enable the routine collection of outcomes that matter to coronary revascularisation patients, rather than focus on narrow aspects of disease or functioning. The CROQ is not yet appropriate for use in clinical practice at the individual patient level as it was developed using psychometric tests for group level measurement and more rigorous measurement standards need to be met for this application.[33]

Author affiliations
[1]The BMJ, London, UK
[2]London School of Hygiene and Tropical Medicine, London, UK
[3]Miles & Green Associates, Oxford, UK
[4]Institute of Cardiovascular Medicine and Science, Liverpool Heart and Chest Hospital, Liverpool, UK

Acknowledgements We thank all patients who participated in the NHS Coronary Revascularisation PROMs Pilot and the staff who so generously gave their time voluntarily to help make the pilot work. The participating NHS Trusts who participated in the pilot included: Barts Health NHS Trust, Basildon and Thurrock University Hospitals NHS Foundation Trust (Essex Cardiothoracic Centre), Brompton & Harefield NHS Foundation Trust, Blackpool, Fylde and Wyre Hospitals NHS Foundation Trust, Liverpool Heart & Chest Hospital NHS Foundation Trust, Nottingham University Hospitals NHS Trust, Oxford University Hospitals NHS Foundation Trust, Papworth Hospital NHS Foundation Trust, Sheffield Teaching Hospitals NHS Foundation Trust, St George's NHS Trust, and University Southampton Hospitals NHS Foundation Trust. We thank Dr Andrew Wragg, Mr Peter Bradley and Alison Pottle for their contribution to the working group.

Contributors SS, RM, SG and MJ developed the sampling strategy for the psychometric testing. SG and RM cleaned the data sets and developed the samples for analysis. SS conducted all the psychometric analysis and wrote the first draft of the manuscript. SS, RM, SG, MJ contributed to the writing of the article and approved the final version of the manuscript.

Funding This work was commissioned by the Department of Health (now NHS England).

Competing interests SS developed and validated the CROQ. SS is employed full time by BMJ Publishing Group as a researcher, but is not involved in any publication decisions on manuscripts for any of its journals. A grant was paid by the Department of Health to Liverpool Heart and Chest Hospital to cover the costs of the analytics. RM, SG and SS were compensated for their contributions to the analysis of the NHS Coronary revascularisation PROMs Pilot.

Ethical approval Ethical approval was not required as the data was collected as part of a service evaluation for the NHS. Patients completed a consent form at the time they completed the Q1 questionnaire.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Thompson DR, Yu CM. Quality of life in patients with coronary heart disease-I: assessment tools. *Health Qual Life Outcomes* 2003;1:42.
2. Mackintosh A, Gibbons E, Casanas i Comabella C, *et al. A structured review of patient-reported outcome measures used in elective procedures for coronary revascularisation, 2010*. Patient-reported Outcome Measurement Group, Department of Public Health, University of Oxford, 2010.
3. Schroter S, Lamping DL. Coronary revascularisation outcome questionnaire (CROQ): development and validation of a new, patient based measure of outcome in coronary bypass surgery and angioplasty. *Heart* 2004;90:1460–6.
4. Schroter S, Lamping DL. Responsiveness of the coronary revascularisation outcome questionnaire compared with the SF-36 and Seattle Angina Questionnaire. *Qual Life Res* 2006;15:1069–78.
5. Devlin NJ, Appleby J. Getting the most out of PROMs: putting health outcomes at the heart of NHS decision-making. *Health Econ* 2010. http://www.kingsfund.org.uk/publications/proms.html
6. Department of Health (2008). *Guidance on the routine collection of patient reported outcome measures (PROMs). For the NHS in England 2009/10*. London: Department of Health. https://www.racp.edu.au/docs/default-source/default-document-library/guidance-on-the-routine-collection-of-patient-reported-outcome-measures-(proms)-(pdf-1184-kb)-nhs-(2008).pdf?sfvrsn=4 (accessed 20 Dec 2016).
7. Devlin NJ, Parkin D, Browne J. Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. *Health Econ* 2010;19:886–905.
8. NHS Group, Department of Health. *The NHS Outcomes Framework 2015/16*. London: Department of Health, December 2014. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/385749/NHS_Outcomes_Framework.pdf
9. https://www.england.nhs.uk/statistics/statistical-work-areas/proms/cardiac-update/.html (accessed 16 Dec 2016).
10. EuroQol Group. EuroQol: a new facility for the measurement of health related quality of life. *Health Policy* 1990;16:199–208.

11. Lord F, Novick M. *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley, 1968.
12. Novick M. The axioms and principal results of classical test theory. *J Math Psychol* 1966;3:1–18.
13. Nunnally JC, Bernstein IH. *Psychometric theory*. 3rd edn. New York: McGraw-Hill, 1994.
14. Cappelleri JC, Lundy J, Hays RD. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clin Ther* 2014;36:648–62.
15. Traub R. Classical test theory in historical perspective. *Educational Measurement Issues and Practice* 1997;16:8–14.
16. Petrillo J, Cano SJ, McLeod LD, *et al*. Patient-reported outcomes Using Classical Test Theory, Item Response Theory, and Rasch Measurement Theory to evaluate patient-reported outcome measures: a comparison of worked examples. *Value Health* 2015;18:25–34.
17. Streiner DL, Norman GR. *Health measurement scales: a practical guide to their development and use*. 2nd edn. Oxford: Oxford University Press, 1995.
18. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims, 2009. http://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf
19. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes* 2006;4:79.
20. Schuman H, Presser S. *Questions and answers in attitude surveys*. New York: Academic Press, 1981.
21. Perkins JJ, Sanson-Fisher RW. An examination of self- and telephone-administered modes of administration for the Australian SF-36. *J Clin Epidemiol* 1998;51:969–73.
22. Bowling A, Bond M, Jenkinson C, *et al*. Short Form 36 (SF-36) Health Survey questionnaire: which normative data should be used? Comparisons between the norms provided by the Omnibus Survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *Journal of Public Health Medicine* 1999;21:255–70.
23. Bowling A. Mode of questionnaire administration can have serious effects on data quality. *J Public Health (Oxf)* 2005;27:281–91.
24. Crocker TF, Smith JK, Skevington SM. Family and professionals underestimate quality of life across diverse cultures and health conditions: systematic review. *J Clin Epidemiol* 2015;68:584–95.
25. Spector PE. *Summated rating scale construction: an introduction*. Newbury Park, CA: Sage, 1992.
26. Ware JE, Harris WJ, Gandek B, *et al*. *MAP-R for Windows: multitrait/multi-item analysis program-revised user's guide version 1*. Boston, MA: Health Assessment Lab, 1997.
27. Cohen J. *Statistical power analysis for the behavioural sciences*. Revised edn. New York: Academic Press, 1977.
28. Spertus JA, Winder JA, Dewhurst TA, *et al*. Development and evaluation of the Seattle Angina Questionnaire: a new functional status measure for coronary artery disease. *J Am Coll Cardiol* 1995;25:333–41.
29. Valenti L, Lim L, Heller RF, *et al*. An improved questionnaire for assessing quality of life after acute myocardial infarction. *Qual Life Res* 1996;5:151–61.
30. Ferrans CE, Powers MJ. Quality of Life Index: development and psychometric properties. *ANS Adv Nurs Sci* 1985;8:15–24.
31. Nguyen TH, Han HR, Kim MT, *et al*. An introduction to item response theory for patient-reported outcome measurement. *Patient* 2014;7:23–35.
32. Rasch G. *Probabilistic models for some intelligence and attainment tests*. IL, USA: The University of Chicago Press, 1980.
33. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995;4(4):293–307.