



A systematic review of patient reported outcome measures (PROMs) and quality of life reporting in patients undergoing laparoscopic cholecystectomy

Prita Daliya¹, Elizabeth H. Gemmill², Dileep N. Lobo^{1,3}, Simon L. Parsons^{1,4}

¹Gastrointestinal Surgery, Nottingham Digestive Diseases Centre and National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and University of Nottingham, Queen's Medical Centre, Nottingham NG7 2UH, UK; ²Department of General Surgery, Sherwood Forest Hospitals NHS Trust, King's Mill Hospital, Sutton-in-Ashfield NG17 4JL, UK; ³MRC/ARUK Centre for Musculoskeletal Ageing Research, School of Life Sciences, University of Nottingham, Queen's Medical Centre, Nottingham NG7 2UK, UK; ⁴Trent Oesophago-Gastric Unit, City Hospital Campus, Nottingham University Hospitals NHS Trust, Nottingham NG5 1PB, UK

Contributions: (I) Conception and design: All authors; (II) Administrative support: DN Lobo, SL Parsons; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: P Daliya, EH Gemmill; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Professor Dileep N. Lobo. Professor of Gastrointestinal Surgery, Nottingham Digestive Diseases Centre, Nottingham University Hospitals NHS Trust, Queen's Medical Centre, Nottingham NG7 2UH, UK. Email: dileep.lobo@nottingham.ac.uk.

Abstract: Patient reported outcome measures (PROMs) provide a valuable means of measuring outcomes subjectively from a patient's perspective, facilitating the assessment of service quality across healthcare providers, and assisting patients and clinicians in shared decision making. The primary aim of this systematic review was to critically appraise all historic studies evaluating patient reported quality of life, in adult patients undergoing laparoscopic cholecystectomy for symptomatic gallstones. The secondary aim was to perform a quality assessment of cholecystectomy-specific PROM-validation studies. A literature review was performed in PubMed, Google ScholarTM, the Cochrane Library, Medline, CINAHL, EMBASE and PsychINFO databases up to September 2017. Study characteristics, PROM-specific details and a bias assessment were summarised for non-validation studies. A Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) analysis was performed to assess the methodological quality of identified PROM-validation studies. Fifty one studies were found to evaluate health-related quality of life (HRQoL) after laparoscopic cholecystectomy. Although 94.1% of these studies included PROMs as a primary outcome measure, <20% provided level 1 evidence through randomised controlled trials (RCTs). There was significant variation in the selection and reporting of PROMs, with no studies declaring patient involvement in PROM selection, and 88.2% of studies failing to document the management of missing data points, or non-returned surveys (33.3%). In the 6 PROM-validation studies identified, only 5 psychometric properties were evaluated, the findings of which were limited due to the small number of studies. This systematic review identifies a lack in consistency of study design and PRO reporting in clinical trials. Whilst an increasing number of studies are being performed to evaluate PROs, a lack of adherence to existing PRO administration and reporting guidelines is continuing to negatively affect study quality. We recommend that future clinical trials utilizing PROs should adhere to established comprehensive guidelines as described.

Keywords: Cholecystectomy; patient reported outcomes (PROs); patient reported outcome measures (PROMs); quality of life (QoL); Consensus-based Standards for the selection of health Measurement Instruments (COSMIN)

Submitted Aug 22, 2018. Accepted for publication Nov 22, 2018.

doi: 10.21037/hbsn.2019.03.16

View this article at: <http://dx.doi.org/10.21037/hbsn.2019.03.16>

Introduction

Cholecystectomy is the only definitive treatment for patients with symptomatic gallstones, with laparoscopic cholecystectomy being the current gold standard (1-3). In the UK alone over 60,000 cholecystectomies are performed annually, equivalent to approximately 100 procedures per 100,000 population (1), and more than 200 per 100,000 population in parts of Europe, and North America (1,4,5). Despite the therapeutic benefits of surgery and the potential economic savings in preventing further morbidity from gallstone disease, laparoscopic cholecystectomy is not without risks (2).

Patient reported outcomes (PROs) provide a means of measuring various outcomes such as clinical symptoms, patient satisfaction and health-related quality of life (HRQoL) from a patient's perspective subjectively (6,7). Validated questionnaires or patient reported outcome measures (PROMs) are often used to collect PRO data (8). In the National Health Service (NHS) in England, this process has been adopted as mandatory practice for measuring HRQoL in hip and knee replacement surgery, groin hernia repair and varicose vein surgery since April 2009 (9). In addition to comparing the quality of services across healthcare providers, the collection of PROs can also assist patients and clinicians in clinical decision making; by monitoring illness, and the effectiveness of treatment (9-11).

The primary aim of this systematic review was to identify and critically appraise all historic studies evaluating patient reported HRQoL, in adult patients, undergoing laparoscopic cholecystectomy for symptomatic gallstones.

The secondary aim was to perform a quality assessment of cholecystectomy-specific PROM-validation studies using the Consensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist (12).

Methods

Search strategy

A search for all relevant literature was performed using PubMed, Google™ Scholar, the Cochrane Library, and MEDLINE (Ovid) databases in April 2016, and updated in September 2017 to include CINAHL (EBSCO), EMBASE (Ovid), and PsychINFO (Ovid). The following search criteria modified from those developed by the Oxford PROM Group in 2010 (13), were used to identify relevant studies: (cholelithiasis.mp. or cholecystitis.mp. or

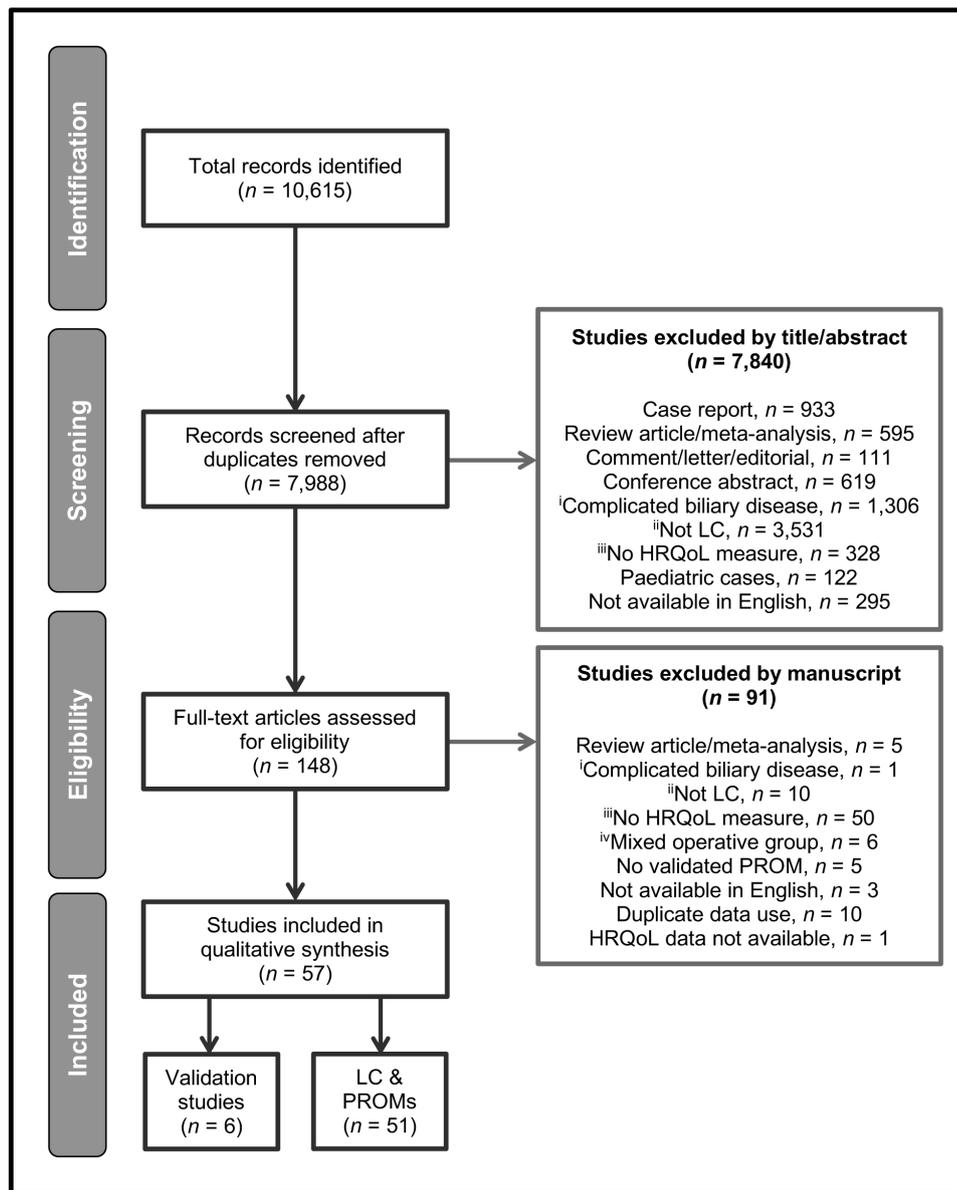
cholecystolithiasis.mp. or gallstone*.mp. or gall stone*.mp. or gallbladder*.mp. or gall bladder*.mp. or biliary colic.mp. or biliary sludge.mp. or cholecystectomy.mp.) and ((HR-PRO or HRPRO or HRQL or HRQoL or QL or QoL or PROM or PRO).ti,ab. or quality of life.mp. or (health index* or health indices or health profile*).ti,ab. or health status.mp. or ((patient or self or child or parent or carer or proxy) adj (appraisal* or appraised or report or reported or reporting or rated or rating* or based or assessed or assessment*)),ti,ab. or ((disability or function or functional or functions or subjective or utility or utilities or wellbeing or well being) adj2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile or profiles or scale or scales or score or scores or status or survey or surveys)).ti,ab.).

The search was performed without date restrictions but was limited to full-text articles. Due to the availability of resources the search was also limited to articles available in the English language, or English translation. Only studies with an adult population over 18 years of age were included. The bibliographies of studies included were also reviewed.

Study selection

Studies identified through the search strategy were assessed for inclusion, initially by title and abstract, and subsequently through full text review (P Daliya and EH Gemmill). Studies were only included where the outcome measure of HRQoL formed either a primary or secondary study aim. Only studies reporting on patients undergoing cholecystectomy for symptomatic gallstones, in which at least one study population underwent a conventional 4-port laparoscopic cholecystectomy (CLC) were included. Studies which reported 'exclusively' on patients with biliary malignancy, and the complications of gallstone disease, such as gallbladder necrosis, perforation, pancreatitis, and choledocholithiasis were excluded due to the potential variability of patient populations and management of these groups.

Validation studies involving either the development or assessment of cholecystectomy-specific PROMs were also included but analysed separately. Review articles such as meta-analyses and systematic reviews were excluded, as were case reports, editorial comments and letters. Duplicate studies and populations were cross-referenced and removed. *Figure 1* demonstrates the preferred reporting items for systematic reviews and meta-analysis (PRISMA) flow diagram (14).



ⁱ**Complicated biliary disease:** gallbladder cancer, gallbladder perforation, gallbladder polyps, gallbladder torsion, biliary dyskinesia, gallstone ileus, cholecysto-X fistula, choledochal cyst, bile duct stones, bile duct injury, biliary vessel aneurysm/pseudo-aneurysm, acute pancreatitis, pancreatic tumour, hepatocellular carcinoma, cholecystostomy, biliary bypass procedure.

ⁱⁱ**Not LC:** no Laparoscopic Cholecystectomy (LC) group.

ⁱⁱⁱ**No HRQoL measure:** Health Related Quality of Life (HRQoL) measurement not taken or described.

^{iv}**Mixed operative group:** Disorders and conditions requiring surgery other than cholecystectomy, e.g., inguinal hernia repair, bowel resection, venous surgery, etc.

Figure 1 Preferred reporting items for systematic reviews and meta-analysis [PRISMA (14)] flow diagram.

Data extraction

Two independent reviewers (P Daliya and EH Gemmill) extracted data from the included studies, with discrepancies resolved by a third and fourth (DN Lobo and SL Parsons). Data were collected on the details of the publication (author names, year of publication, level of evidence and study type, number of centres involved, and country), patient characteristics within each study (sample size, diagnoses, mean age, and gender), and PROM-specific details (PRO instruments used, PRO concepts and scoring methodology, and survey distribution, response, and follow-up). An assessment of bias was performed on all non-validation studies utilising the revised Cochrane risk-of-bias tool for randomised trials (RoB 2.0) (15), and the Risk Of Bias In Non-randomised Studies – of Interventions (ROBINS-I) assessment tool (16), as appropriate.

Quality assessment of cholecystectomy-specific validation studies

The assessment of the quality of PROM-validation studies was performed using the COSMIN checklist; a critical appraisal tool which was devised as part of a Delphi study to help evaluate the methodological quality of studies on PROs (12,17). The checklist uses a standardised descriptive framework to assess each of 9 measurement properties (internal consistency, reliability, measurement error, content validity, structural validity, hypothesis testing, cross-cultural validity, criterion validity, responsiveness) against quality markers. Each measurement property, where relevant, was assessed by completing between 1–18 items on the checklist. A 4-point scoring system (“poor”, “fair”, “good”, and “excellent”) specifically designed by COSMIN for systematic reviews of measurement properties was used to grade each item (12). An overall score for each measurement property was summarised on a “worst score counts” basis, i.e., where a score of “good” or “excellent” was deemed as evidence in support of adequate methodological quality for that study and “poor” or “fair” as inadequate methodological quality (12).

Registration of review

The study proposal was registered (Reg. No. CRD42016048211) with the PROSPERO database (www.crd.york.ac.uk/prospere). This was amended subsequently to include the additional databases used, and the specifications required to complete a

COSMIN analysis on PROM-validation studies.

Results

A total of 10,615 articles were identified and screened by title and abstract review. Of these, 148 remaining articles underwent full text review for eligibility. Details on the use of a PRO questionnaire were frequently found to be lacking, or specifics on the study population such as diagnoses, or intervention were also not defined in some cases. Further details on study exclusion are as described in the PRISMA flow diagram (*Figure 1*) (14).

A total of 57 studies were identified as utilising PROMs in patients undergoing laparoscopic cholecystectomy, of which 6 of these were identified as validation studies researching the psychometric properties for PROMs in patients undergoing laparoscopic cholecystectomy (18-23).

Study quality

Of the 51 non-validation (24-35) studies (36-50) identified (51-74), the majority were performed in the last decade (62.7%), in Europe (60.8%), and as single centre studies (60.8%). Almost 20% provided level 1 evidence through randomised controlled trials (RCTs), but the majority were conducted as either prospective cohort or case control studies. All included trials specified the inclusion of patients with symptomatic gallstone disease, although further analyses identified significant heterogeneity in this definition which also included choledocholithiasis, pancreatitis, biliary dyskinesia, and incidental biliary tumours (*Table 1*).

There was significant variation in the selection of PROMs as reflected by the differing study outcomes, however the 36-item Short Form survey (SF-36) generic measure, European Quality of Life Five Dimensions Questionnaire (EQ5D) utility measure, gastrointestinal quality of life index (GIQLI) disease-specific measure, and visual analogue pain scores (VAPS), featured the most frequently. Study samples in 18 studies (35.3%) were found to be ≤ 100 , although these ranged from between 31 to 100 patients, with 5 studies describing a population of < 60 patients (25,35,49,51,58).

Risk of bias assessment

A risk of bias assessment demonstrated very few studies with a consistently low risk of bias across all domains. Although

Table 1 Summary of study characteristics (characteristics of each of the 51 non-validation studies)

| Characteristics | n | % |
|---|----|------|
| Year of publication | | |
| 1991–2000 | 6 | 11.8 |
| 2001–2010 | 13 | 25.5 |
| 2011–2017 | 32 | 62.7 |
| Origin of study | | |
| Europe | 31 | 60.8 |
| Africa | 1 | 2.0 |
| Asia | 9 | 17.6 |
| North America | 8 | 15.7 |
| Intercontinental | 2 | 3.9 |
| Level of evidence & study type | | |
| Level 1: randomized controlled trials | 10 | 19.6 |
| Level 2: prospective cohort studies | 40 | 78.4 |
| Level 3: case control studies | 1 | 2.0 |
| Number of study sites involved | | |
| 1 | 31 | 60.8 |
| 2 | 8 | 15.7 |
| 3 | 4 | 7.8 |
| 5 | 3 | 5.9 |
| >5 | 5 | 9.8 |
| Diagnoses* | | |
| Cholelithiasis | 38 | 74.5 |
| Acute and chronic cholecystitis | 18 | 35.3 |
| Biliary polyps | 9 | 17.6 |
| Biliary dyskinesia | 5 | 9.8 |
| Choledocholithiasis | 8 | 15.7 |
| Gallstone pancreatitis | 9 | 17.6 |
| Gallbladder tumour | 2 | 3.9 |
| Not specified (symptomatic gallstone disease) | 10 | 19.6 |
| Sample size | | |
| 1–100 | 18 | 35.3 |
| 101–200 | 15 | 29.4 |
| 201–300 | 8 | 15.7 |
| 301–400 | 4 | 7.8 |

Table 1 (continued)**Table 1** (continued)

| Characteristics | n | % |
|--|----|------|
| 401–500 | 2 | 3.9 |
| >500 | 4 | 7.8 |
| PRO as an outcome measure | | |
| Primary outcome | 39 | 76.5 |
| Secondary outcome | 3 | 5.9 |
| Primary & secondary outcome | 9 | 17.6 |
| Number of PROs measured* | | |
| 1 | 18 | 35.3 |
| 2 | 17 | 33.3 |
| >2 | 16 | 31.4 |
| When PRO was measured | | |
| Pre- and post-operatively | 40 | 78.4 |
| Post-operatively only | 10 | 19.6 |
| Not specified | 1 | 2.0 |
| PRO instruments used* | | |
| HRQoL measure | | |
| EuroQoL EQ5D | 8 | 15.7 |
| EQVAS | 3 | 5.9 |
| Short form 8 | 2 | 3.9 |
| Short form 12 | 4 | 7.8 |
| Short form 36 | 17 | 33.3 |
| NHPQ | 6 | 11.8 |
| PESQ | 1 | 2.0 |
| Hospital anxiety and depression scale | 1 | 2.0 |
| PGWB | 2 | 3.9 |
| MOS-24 | 1 | 2.0 |
| PROMIS-10 | 1 | 2.0 |
| Linear analog self-assessment | 1 | 2.0 |
| SFD-LC | 1 | 2.0 |
| State-Trait inventory | 1 | 2.0 |
| Health index | 1 | 2.0 |
| SOMS | 2 | 3.9 |
| Gastrointestinal quality of life index | 17 | 33.3 |
| CSQ | 1 | 2.0 |
| Gastrointestinal symptom survey | 1 | 2.0 |

Table 1 (continued)

Table 1 (continued)

| Characteristics | n | % |
|---------------------------------|----|------|
| Gallstone symptom checklist | 1 | 2.0 |
| Visual analogue scale | 2 | 3.9 |
| Cosmesis & body image measure | | |
| Body image scale | 1 | 2.0 |
| Body image questionnaire | 4 | 7.8 |
| Photograph series questionnaire | 3 | 5.9 |
| POSAS | 1 | 2.0 |
| Likert scale | 1 | 2.0 |
| Visual analogue scale | 2 | 3.9 |
| Pain measure | | |
| Visual analogue pain scale | 15 | 29.4 |
| McGill pain questionnaire | 3 | 5.9 |
| Satisfaction measure | | |
| Likert scale | 1 | 2.0 |
| Visual analogue scale | 1 | 2.0 |
| Distribution of PROs* | | |
| Clinic | 23 | 45.1 |
| Post | 27 | 52.9 |
| Ward | 10 | 19.6 |
| Phone | 7 | 13.7 |
| Email | 1 | 2.0 |
| Researcher | 6 | 11.8 |
| Not specified | 10 | 19.6 |
| Post-operative PRO return rate | | |
| 80–100% | 23 | 45.1 |
| 60–79% | 9 | 17.6 |
| <60% | 4 | 7.8 |
| Not specified | 15 | 29.4 |

*, may be ≥ 1 per study. EuroQoL EQ5D, European Quality of Life Five Dimensions Questionnaire; EQVAS, European Quality of Life Visual Analogue Scale; NHPQ, Nottingham Health Profile Questionnaire; PESQ, Patients' Experience of Surgery Questionnaire; PGWB, Psychological General Well-Being index; MOS-24, Medical Outcomes Study 24-item Short Form Health Survey; PROMIS-10, Patient-Reported Outcomes Measures Information System-10; SFD-LC, Symptom, Frequency and Distress questionnaire-LC version; SOMS, Surgical Outcomes Measurement System; CSQ, Otago Gallstones Condition Specific Questionnaire; POSAS, Patient and Observer Scar Assessment Scale.

randomisation was performed well to minimise selection bias in the majority of studies, blinding was performed quite poorly overall. Where a number of studies employed special dressings to blind patient participants against intervention identification, some comparative outcomes were unable to be realistically blinded against due to the specific outcomes studied (27,28,33,42,44,45,51,71). These included the comparison of inpatient and outpatient cholecystectomy (28,34,42,44), and the measure of cosmesis (33,39,45,51). Three studies were underpowered, having failed to recruit sufficient participants (27,30,48), and 12 presented incomplete data having either excluded surveys with missing responses or discounted those lost to follow-up (30,31,34,36,40,46,47,50,53,57,65,68) (Tables 2,3).

PROMs

The majority of studies (94.1%) included PROs as a primary outcome measure, and over 60% of studies measured more than two PRO concepts. These included HRQoL, cosmesis and body image, post-operative pain, sexual function and patient satisfaction (Table 4). Forty studies (78.4%) were set up with the intent to compare two or more different operative techniques for cholecystectomy. All studies were performed for research purposes with no involvement of patient groups to aid PROM selection. Profile scores rather than single indicator or index numbers were used to describe PROs in the majority of studies, with only 25.5% of studies using both generic and disease-specific PROMs, and only 60.8% of studies used PROMs which were validated with demonstrable evidence of this. A significant proportion of studies (88.2%) did not discuss the management of missing responses within surveys, 21.6% did not consider baseline or pre-operative PROM scores for their population, and 33.3% were not clear about their survey return rate. Full study characteristics are available in the Table S1.

Validation studies

Of the 6 PROM-validation studies identified, 4 reported on the gastrointestinal quality of life index (GIQLI) (18-21), whereas one study reported on the Otago Gallstones Condition Specific Questionnaire (CSQ) (22), and one on the Gallstone Impact Checklist (GIC) (23). These studies included original validation studies (18,22,23), in addition to translations in to other languages (19-21).

Table 2 Risk of bias assessment [risk of bias assessment using the Revised Cochrane risk-of-bias for randomised trials (RoB 2.0)]

| Study | Random sequence generation (selection bias) | Allocation concealment (selection bias) | Blinding of participants and personnel (performance bias) | Blinding of outcome assessment (detection bias) | Incomplete outcome data (attrition bias) | Selective reporting (reporting bias) | Other bias | Notes |
|------------------------|---|---|---|---|--|--------------------------------------|------------|--------------------------------|
| Abd Ellatif, 2013 (24) | + | + | + | + | + | + | - | Single centre |
| Ainslie, 2003 (25) | + | + | + | + | + | + | - | Single centre |
| Aspinen, 2017 (26) | + | + | + | ? | + | + | - | No power calculation |
| Barkun, 1992 (27) | + | ? | - | + | - | ? | - | Underpowered |
| Barthelsson, 2006 (28) | + | - | - | + | + | ? | - | Single centre |
| Bingener, 2015 (29) | + | + | + | + | ? | ? | - | Single centre |
| Brown, 2013 (32) | ? | + | + | + | + | + | - | Single centre |
| Bucher, 2011 (33) | - | - | - | ? | + | + | - | Mixed diagnoses. Single centre |
| Harju, 2007 (38) | ? | ? | ? | ? | - | ? | - | Single centre |
| Johansson, 2006 (42) | + | - | - | ? | ? | + | - | Single centre |
| Keus, 2008 (43) | + | + | + | + | + | + | - | Single centre |
| Keulemans, 1998 (44) | + | + | - | ? | + | + | - | Single centre |
| Kudsi, 2017 (45) | - | - | - | - | - | + | - | Mixed diagnoses |
| Leung, 2012 (48) | - | + | ? | + | ? | ? | - | Underpowered. Mixed diagnoses |
| Liu, 2016 (50) | + | + | ? | + | - | ? | - | Lost to follow-up |
| Ma, 2011 (51) | - | ? | - | - | + | ? | - | Mixed diagnoses |
| Nilsson, 2004 (55) | + | + | ? | ? | - | + | - | - |

Table 2 (continued)

Table 2 (continued)

| Study | Random sequence generation (selection bias) | Allocation concealment (selection bias) | Blinding of participants and personnel (performance bias) | Blinding of outcome assessment (detection bias) | Incomplete outcome data (attrition bias) | Selective reporting (reporting bias) | Other bias | Notes |
|------------------------|---|---|---|---|--|--------------------------------------|------------|---|
| Phillips, 2012 (57) | - | + | + | ? | - | ? | - | Groups not equivalent. Lost to follow-up. Mixed diagnoses |
| Rosenmüller, 2017 (61) | - | ? | ? | ? | + | ? | - | Groups not equivalent. Mixed diagnoses |
| Saad, 2013 (63) | + | + | + | + | + | + | - | Single centre |
| Squirrel, 1998 (66) | + | + | + | ? | ? | ? | - | Single centre |
| Sulu, 2015 (67) | + | ? | ? | ? | ? | ? | - | Single centre |
| Vetthus, 2004 (71) | + | - | - | ? | + | ? | - | - |

Risk of bias assessment: + = low; ? = unclear; - = high.

COSMIN analysis

The commonest measurement properties analysed were internal consistency and reliability (all 6 PROM-validation studies), and responsiveness (5 of 6 studies). Only 2 studies scored either “good” or “excellent” for internal consistency, describing adequate methodological quality (20,21), whereas the other 4 studies rated as either “fair” or “poor”, describing inadequate methodological quality. The summary scores for each measurement property, for each study are shown in Table 5. No studies performed an assessment of “measurement error”, “hypotheses testing”, or “criterion validity”. The methodological qualities assessed for each study are summarised in Table S2.

Due to the limited number of PROM-validation studies identified, the quality of the measurement instruments identified was not assessed against the “criteria for good measurement properties” as recommended by the COSMIN guidelines (76) and, therefore, preclude recommendation of a specific PROM for use in laparoscopic cholecystectomy

Discussion

PROM selection

A recent systematic review of RCTs evaluating PROs after cholecystectomy (77) utilized the International Society of Quality of Life Research (ISOQOL) checklist to assess the quality of reporting in their evaluated studies. The authors demonstrated that, despite the availability of the ISOQOL checklist since 2013, the majority of studies did not adhere to guidelines, and demonstrated high bias and poor quality reporting of PROs (77).

In contrast, we analyzed all clinical trials evaluating HRQoL after laparoscopic cholecystectomy so as not to exclude the majority of clinical studies (>80%) which were non-RCTs. We therefore used the amended checklist as described by Patrick and Erikson (75) in the Cochrane Handbook, to describe and assess the identified studies. Much like Mueck *et al.* (77) the present review also demonstrated significant variability in PRO reporting. Across the clinical trials included, a wide variety of concepts were evaluated in addition to HRQoL, via a number of different PRO instruments (Table 1). This variation reflects the lack of specific recommendations in PROM selection in patients undergoing laparoscopic cholecystectomy, and the variation in study rationale which in itself can impact PROM selection.

Each study seemingly selected PRO instruments

Table 3 Risk of bias assessment [risk of bias assessment using the Risk Of Bias In Non-randomized Studies-of Interventions (ROBINS-I) assessment tool]

| Study | Confounding | Selection bias | Bias in classification of interventions | Bias due to deviations from intended interventions | Incomplete outcome data (attrition bias) | Blinding of outcome assessment (detection bias) | Selective reporting (reporting bias) | Other bias | Notes |
|---------------------|-------------|----------------|---|--|--|---|--------------------------------------|------------|---|
| Bitzer, 2008 (30) | ? | - | + | + | - | - | ? | - | Underpowered. Incomplete data |
| Borchert, 2012 (31) | + | + | + | + | - | ? | ? | - | Incomplete data. Single centre |
| Burney, 2002 (34) | + | + | + | + | - | - | ? | - | Lost to follow-up. Single centre |
| Chen, 2005 (35) | + | - | + | + | ? | - | - | - | Single centre |
| Cleary, 1995 (36) | + | ? | ? | ? | - | - | ? | - | Lost to follow-up. Single centre |
| Finan, 2006 (37) | + | + | + | ? | + | + | + | - | Single centre |
| Hauters, 2013 (39) | + | - | + | + | + | ? | ? | - | - |
| Howie, 2017 (40) | ? | + | ? | ? | - | ? | ? | - | Lost to follow-up |
| Hsueh, 2011 (41) | + | + | + | ? | ? | ? | ? | - | - |
| Kum, 1996 (46) | ? | - | + | ? | - | - | ? | - | Incomplete data. Groups not equivalent. Single centre |
| Lamberts, 2015 (47) | + | + | + | + | - | - | - | - | Groups not equivalent. Lost to follow-up |
| Lirici, 2011 (49) | + | + | + | + | + | + | ? | - | - |
| Matovic, 2012 (52) | + | ? | + | ? | ? | ? | ? | - | Single centre |
| McLean, 2017 (53) | ? | ? | + | + | - | + | ? | - | Groups not equivalent. Lost to follow-up. Single centre |
| Mentes, 2001 (54) | ? | + | + | + | ? | + | ? | - | Single centre |

Table 3 (continued)

Table 3 (continued)

| Study | Confounding | Selection bias | Bias in classification of interventions | Bias due to deviations from intended interventions | Incomplete outcome data (attrition bias) | Blinding of outcome assessment (detection bias) | Selective reporting (reporting bias) | Other bias | Notes |
|---------------------------|-------------|----------------|---|--|--|---|--------------------------------------|------------|---|
| Pälsson, 2011 (56) | - | ? | ? | ? | - | ? | ? | - | Mixed diagnoses |
| Plaisier, 1995 (58) | + | ? | + | ? | + | + | + | - | Single centre |
| Plannells Roig, 2013 (59) | - | ? | - | ? | + | ? | ? | - | Mixed diagnoses. Mixed interventions. Single centre |
| Reibetanz, 2013 (60) | + | ? | + | + | + | ? | ? | - | Single centre |
| Rydbeck, 2015 (62) | - | ? | ? | ? | ? | ? | ? | - | Mixed diagnoses. Mixed interventions |
| Sadati, 2016 (64) | ? | ? | + | + | ? | ? | ? | - | - |
| Shi, 2011 (65) | + | + | + | + | - | ? | - | - | Lost to follow-up |
| Tani, 2015 (68) | + | ? | + | + | - | ? | ? | - | Lost to follow-up. Single centre |
| Teubner, 2016 (69) | ? | - | + | ? | ? | ? | ? | - | - |
| Topcu, 2003 (70) | - | - | + | + | + | ? | ? | - | Groups not equivalent. Single centre. Mixed interventions |
| Wagner, 2013 (72) | - | ? | + | ? | ? | ? | ? | - | Mixed diagnoses. Single centre |
| Wanjura, 2016 (73) | - | - | - | ? | ? | ? | - | - | Mixed diagnoses. Single centre |
| Zapf, 2013 (74) | - | ? | + | + | ? | ? | - | - | Mixed diagnoses |

Risk of bias assessment: + = low; ? = unclear; - = high.

Table 4 Summary of patient reported outcomes measures (PROMs) used in the clinical trials (n=51)

| Variable | n | % |
|---|----|------|
| What were the PROs measuring? | | |
| What concepts were the PROs used in the study measuring?* | | |
| Health-related quality of life (HRQoL) | 51 | 100 |
| Cosmesis & body image | 12 | 23.5 |
| Pain | 18 | 35.3 |
| Sexual function | 1 | 2.0 |
| Satisfaction | 5 | 9.8 |
| What rationale (if any) for the selection of concepts or constructs did the authors provide? | | |
| To compare PROs between 2 or more cholecystectomy types | 40 | 78.4 |
| To assess validity of a monitoring system | 1 | 2.0 |
| To assist patient selection between 2 or more cholecystectomy types | 1 | 2.0 |
| To correlate clinical findings predictive with a better quality of life (QoL) | 8 | 15.7 |
| To incorporate QoL into routine clinical practice | 1 | 2.0 |
| Were patients involved in the selection of outcomes measured by the PROs? | | |
| No | 51 | 100 |
| Omissions | | |
| Were there any important aspects of health or QoL that were omitted in this study?* | | |
| Validated disease-specific HRQoL instrument not considered | 32 | 62.7 |
| Validated generic HRQoL instrument not considered | 7 | 13.7 |
| Baseline or pre-operative QoL score not considered | 11 | 21.6 |
| Management of missing HRQoL responses not discussed | 45 | 88.2 |
| Percentage of surveys returned not discussed or unclear | 17 | 33.3 |
| If randomized trials and other studies measured PROs, what were the instruments' measurement strategies? | | |
| Did investigators use HRQoL instruments that yield a single indicator or index number, or a profile score?* | | |
| Profile scores | 45 | 88.2 |
| Single indicator/index number (overall score) | 27 | 52.9 |
| If investigators measured PROs, did they use specific or generic measures, or both? | | |
| Specific | 9 | 17.6 |
| Generic | 29 | 56.9 |
| Both | 13 | 25.5 |
| Who exactly completed the instruments? | | |
| Patient | 50 | 98.0 |
| Patient & interviewer | 1 | 2.0 |

Table 4 (continued)

Table 4 (continued)

| Variable | n | % |
|--|----|------|
| Did the instruments work in the way they were supposed to work? | | |
| Had the instruments used been validated previously, and was evidence of prior validation for use in this population presented?* | | |
| Validated, but not evidenced in the study | 28 | 54.9 |
| Validated, and evidenced in the study | 31 | 60.8 |
| Not validated or evidenced | 8 | 15.7 |
| Were the instruments re-validated in this study? | | |
| No | 49 | 96.1 |
| Yes | 2 | 3.9 |
| Did the instruments work in the way they were supposed to work—ability to measure change? | | |
| Are the PROs able to detect change in patient status, even if those changes are small? | | |
| Yes | 51 | 100 |
| Can you make the magnitude of effect understandable to readers? | | |
| Can you provide an estimate of the difference in patients achieving a threshold of function or improvement, and the associated number needed to treat? | | |
| Possible from the data presented | 20 | 39.2 |
| Not possible from the data presented | 31 | 60.8 |

*, may be ≥1 per study. Checklist amended from that used by Patrick and Erikson (75) in the Cochrane Handbook. PROs, patient reported outcomes; QoL, quality of life.

Table 5 COSMIN (Consensus-based Standards for the selection of health Measurement INstruments) analysis (methodological quality of each study per Patient Reported Outcome Measure questionnaire per measurement property)

| Study | Internal consistency | Reliability | Measurement error | Content validity | Structural validity | Hypotheses testing | Cross-cultural validity | Criterion validity | Responsiveness |
|--|----------------------|-------------|-------------------|------------------|---------------------|--------------------|-------------------------|--------------------|----------------|
| GIQLI: Gastro-Intestinal Quality of Life Index | | | | | | | | | |
| Eypasch, 1995 (18) | Fair | Poor | – | Fair | – | – | – | – | Poor |
| Lien, 2007 (19) | Poor | Poor | – | – | – | – | Fair | – | Poor |
| Quintana, 2001 (20) | Excellent | Poor | – | – | – | – | Poor | – | Poor |
| Sandblom, 2009 (21) | Good | Fair | – | – | – | – | Poor | – | Poor |
| CSQ: Otago Gallstones Condition Specific Questionnaire | | | | | | | | | |
| Chen, 2006 (22) | Poor | Poor | – | Fair | – | – | – | – | – |
| GIC: Gallstone Impact Checklist | | | | | | | | | |
| Russell, 1996 (23) | Poor | Poor | – | Fair | – | – | – | – | Poor |

based on the relevance to primary or secondary outcomes. However, despite the availability of guidance documentation on the use of PROs in clinical trials (11,75), only 25% of the studies reviewed measured both generic and condition-specific PROs. Justification on the rationale for selection was also varied, with documentation in only 16 papers. These reasons included the following: due to the availability of a standardized comparative reference population (30,56,59,73), pre-existing validation within the same or similar cohort (32,40,47,54,56,58,65,71), easier survey application or user friendliness (26,34), adherence to recommended guidelines (although these were not specified) (43), to aid the calculation of specific outcomes which are dependent on a specific type of PROM, i.e., quality-adjusted life year (QALY) (61), or prior knowledge of the psychometric quality of the chosen instrument (28).

PROM dissemination

To our knowledge the completion of PRO surveys in all specified studies was either by the patients themselves or with the aid of a dedicated researcher. These were administered to patients in person, by post, or over the phone via traditional paper surveys. None of the studies included described the use of digital, electronic, or automated PRO mediums, despite recent advances in technology. Many alternative modalities are now available including web-based patient surveys, tablet-based applications, or voice activated phone surveys to name a few (78).

Whilst these more modern methods have the potential to increase the efficiency of data collection, reduce transcription errors caused by data entry, aid data analysis, and reduce missing data points within surveys, they also have some significant limitations. Licensing costs for validated surveys can be significant given the importance of data security, and users must be technologically adept or receive appropriate training (78).

All studies collected PRO data prospectively for research purposes although 4 (40,56,61,62) of the 8 Swedish studies also utilized their national registry [GallRiks (79)] to aid data collection as standard practice, which was also performed prospectively. This is significant given that retrospective data collection is more likely to add bias due to poor recollection and a potential increase in data gaps (78).

PRO analysis

In trials where both profile and index score calculations were possible, there was no explanation given when only one measurement strategy was used. Although profile scores can provide useful information on multiple PRO domains such as physical (pain, mobility, activity) and psychological (mood, energy, anxiety or depression) functioning, they are not always possible nor do they provide additional benefit when compared with index or indicator scores in some studies (75). These overall scores can provide sufficient information to demonstrate a change in HRQoL and can be particularly useful when PROs are used as markers for other outcomes such as cost effectiveness information or QALY to assess service quality (75).

Although the majority of studies commented on the return rate of surveys, which demonstrated good overall patient participation and low attrition, very few studies commented on the management of incomplete returned HRQoL surveys. This is significant as the imputation of results into missing data points or conversely the extraction of incomplete surveys can introduce bias (8,78). Similarly the lack of pre-operative or baseline population values prevent a calculation in change from baseline; a concept useful in demonstrating unbiased improvements or deterioration from the population norm (80).

Review of methodological quality

The 2011 review by the Oxford PROM Group appraised 7 PRO instruments on methodological quality, and performed an expert panel review on their suitability as clinical care evaluation tools (11). This review recommended one of two generic health measures (SF36), one of one preference-based measures (EQ5D), and one of four condition-specific measures (CSQ) as PRO assessment tools in patients undergoing cholecystectomy. Interestingly, these recommended PROs were noted to have little or no evidence of good methodological quality on assessment of their psychometric properties (reproducibility, internal consistency, content validity, construct validity, responsiveness, interpretability, and the presence of floor to ceiling effects). Similarly, our COSMIN review of the 6 PROM-validation studies identified, also demonstrated fair to poor methodological quality in the majority of the psychometric properties evaluated (internal consistency, reliability, content validity, cross-cultural validity, and

responsiveness). Unfortunately, the identification of such few validation studies of poor quality obviated the possibility of commenting on the quality of the identified PRO instruments. This has a significant bearing on recommending PRO instruments as guidelines suggest that studies of poor quality provide little value (12).

Psychometric properties in detail

Internal consistency was analysed in all 6 PROM-validation studies. Despite all studies achieving a Cronbach's alpha ≥ 0.7 for their global rating score, COSMIN analysis demonstrated poor internal consistency. This was accounted for by studies failing to describe their management of missing data points (18), and the use of inadequate sample sizes (19,22,23). Further inspection also demonstrated that where global rating scales achieved an optimum acceptable value ≥ 0.7 for Cronbach's alpha (7,11,76), a measure of scale reliability, individual dimension scores were found to be < 0.7 in some instances (19,23) demonstrating poor inter-correlation within scales.

Studies deemed to have poor reliability (a measure of scale stability) had small sub-group sample sizes (18), and inadequate intervals (18) [such as 48 hours, instead of the recommended minimum of 2 weeks (81)]. The performance of significant interventions between test-retest readings [surgical management (20,23)], a change in environment [ward-based to clinic room (22)], or a move from researcher-led surveys, to postal surveys (22) also affected the measure of scale reliability (76).

The three studies (18,22,23) which measured content validity performed well overall, only missing out on a score of excellent due to minor methodological flaws in study design, such as non-reporting of missing data, or a lack of sub-group demographics to detail the constitution of the expert review panel.

Three studies involved a translation of the GIQLI questionnaire (19-21). Two were translated from the original German GIQLI to either Spanish (20) or Swedish (21), and one was translated from the original English GIQLI to Mandarin Chinese (19). Two studies commented on translation alone and did not meet the full criteria for cross-cultural validity (19,21). Information on the expertise of translators was limited to language expertise alone in all studies. There was no description on the expertise of translators with respect to the disease process studied, or the construct measured. No mention was made of whether translators worked independently, and all studies performed

the minimum requisite of one forward and backward translation, using a minimum of 2 translators. The translation studies did not describe any pre-test process (19,21), although in the one study that analysed full cross-cultural validity, minimal information was provided of the study sub-group, reducing the overall methodological score to "poor" (20).

All studies evaluated scored poor for responsiveness due to the absence of detail on study hypotheses (18-21,23). This was because none of the studies had commented on or quantified the expected direction or magnitude of study outcomes a-priori (76).

Strengths and limitations

Although other assessments of methodological quality are available, the COSMIN analysis is to our knowledge the most standardized method of assessment of PROM-validation studies given the stringent criteria and associated guidelines.

Due to our limitation of resources the exclusion of studies performed in languages other than English may have prevented the identification of some PRO and PROM-validation studies.

Conclusions

This review of PRO studies assessing HRQoL and PROM-validation studies in patients undergoing laparoscopic cholecystectomy identifies a lack in consistency of study design and PRO reporting in clinical trials. Whilst an increasing number of studies are being performed to evaluate PROs, a lack of adherence to existing PRO administration and reporting guidelines is continuing to negatively affect study quality. We recommend future clinical trials utilizing PROs should adhere to established comprehensive guidelines as described in the CONSORT (Consolidated Standards of Reporting Trials) PRO extension (6), and the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) PRO extension (80). Researchers should aim to re-validate PRO instruments in their study population (75,80) and, therefore, ensure selected PROMs have good methodological quality (76).

Acknowledgments

Funding: This work was supported by the Medical Research Council [grant number MR/K00414X/1]; and Arthritis

Research UK [grant number 19891]. Prita Daliya is a recipient of a Research Fellowship funded by the Royal College of Surgeons of England and EIDO Healthcare Limited.

Footnote

Conflicts of Interest: SL Parsons is a company director for EIDO Healthcare Limited. The other authors have no conflicts of interest to declare.

References

1. Aerts R, Penninckx F. The burden of gallstone disease in Europe. *Aliment Pharmacol Ther* 2003;18 Suppl 3:49-53.
2. Gurusamy KS, Davidson BR. Gallstones. *BMJ* 2014;348:g2669.
3. Association of Upper Gastrointestinal Surgeons of Great Britain and Ireland. Commissioning Guide: Gallstone Disease. London: Royal College of Surgeons of England, 2016.
4. Sinha S, Hofman D, Stoker DL, et al. Epidemiological study of provision of cholecystectomy in England from 2000 to 2009: retrospective analysis of Hospital Episode Statistics. *Surg Endosc* 2013;27:162-75.
5. Shaffer EA. Gallstone disease: epidemiology of gallbladder stone disease. *Best Pract Res Clin Gastroenterol* 2006;20:981-96.
6. Calvert M, Blazeby J, Altman DG, et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA* 2013;309:814-22.
7. Deshpande PR, Rajan S, Sudeepthi BL, et al. Patient-reported outcomes: a new era in clinical research. *Perspect Clin Res* 2011;2:137-44.
8. Kyte D, Ives J, Draper H, et al. Current practices in patient-reported outcome (PRO) data collection in clinical trials: a cross-sectional survey of UK trial staff and management. *BMJ Open* 2016;6:e012281.
9. Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;346:f167.
10. Department of Health. Guidance on the routine collection of Patient Reported Outcome Measures (PROMs). London: Department of Health, 2008:1-28.
11. Casanas i Comabella C, Gibbons E, Fitzpatrick R. A structured review of patient-reported outcome measures for patients undergoing cholecystectomy. In: Health Do, editor. Department of Public Health, University of Oxford: Patient-Reported Outcome Measurement Group, 2011:1-40.
12. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651-7.
13. MacIntosh A, Casanas i Comabella C, Hadi M, et al. PROM Group Construct & Instrument Type Filters February 2010. The Oxford PROM Group. 2010. Available online: <http://www.cosmin.nl/images/upload/files/PROM%20Gp%20filtersOCTOBER%202010FINAL.pdf>. 2016.
14. Moher D, Liberati A, Tetzlaff J, et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Open Med* 2009;3:e123-30.
15. Higgins JPT, Sterne JAC, Savović J, et al. A revised tool for assessing risk of bias in randomized trials. In: Chandler J, McKenzie J, Boutron I, et al. editors. *Cochrane Methods*. *Cochrane Database Syst Rev* 2016;10(Suppl 1).
16. Sterne JAC, Hernán MA, Reeves BC, et al. ROBINS-I; a tool for assessing risk of bias in non-randomized studies of interventions. *BMJ* 2016;355:i4919.
17. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539-49.
18. Eypasch E, Williams JI, Wood-Dauphinee S, et al. Gastrointestinal Quality of Life Index: development, validation and application of a new instrument. *Br J Surg* 1995;82:216-22.
19. Lien HH, Huang CC, Wang PC, et al. Validation assessment of the Chinese (Taiwan) version of the Gastrointestinal Quality of Life Index for patients with symptomatic gallstone disease. *J Laparoendosc Adv Surg Tech A* 2007;17:429-34.
20. Quintana JM, Cabriada J, Lopez de Tejada I, et al. Translation and validation of the gastrointestinal Quality of Life Index (GIQLI). *Revista Espanola de Enfermedades Digestivas* 2001;93:693-706.
21. Sandblom G, Videhult P, Karlson BM, et al. Validation of Gastrointestinal Quality of Life Index in Swedish for assessing the impact of gallstones on health-related quality of life. *Value Health* 2009;12:181-4.
22. Chen TY, Landmann MG, Potter JC, et al. Questionnaire to aid priority and outcomes assessment in gallstone disease. *ANZ J Surg* 2006;76:569-74.
23. Russell ML, Preshaw RM, Brant RF, et al. Disease-specific quality of life: the Gallstone Impact Checklist. *Clin Invest Med* 1996;19:453-60.

24. Abd Ellatif ME, Askar WA, Abbas AE, et al. Quality-of-life measures after single-access versus conventional laparoscopic cholecystectomy: a prospective randomized study. *Surg Endosc* 2013;27:1896-906.
25. Ainslie WG, Catton JA, Davides D, et al. Micropuncture cholecystectomy vs conventional laparoscopic cholecystectomy: a randomized controlled trial. *Surg Endosc* 2003;17:766-72.
26. Aspinen S, Karkkainen J, Harju J, et al. Improvement in the quality of life following cholecystectomy: a randomized multicenter study of health status (RAND-36) in patients with laparoscopic cholecystectomy versus minilaparotomy cholecystectomy. *Qual Life Res* 2017;26:665-71.
27. Barkun JS, Barkun AN, Sampalis JS, et al. Randomised controlled trial of laparoscopic versus mini cholecystectomy. The McGill Gallstone Treatment Group. *Lancet* 1992;340:1116-9.
28. Barthelsson C, Anderberg B, Ramel S, et al. Outpatient versus inpatient laparoscopic cholecystectomy: a prospective randomized study of symptom occurrence, symptom distress and general state of health during the first post-operative week. *J Eval Clin Pract* 2008;14:577-84.
29. Bingener J, Skaran P, McConico A, et al. A double-blinded randomized trial to compare the effectiveness of minimally invasive procedures using patient-reported outcomes. *J Am Coll Surg* 2015;221:111-21.
30. Bitzer EM, Lorenz C, Nickel S, et al. Assessing patient-reported outcomes of cholecystectomy in short-stay surgery. *Surg Endosc* 2008;22:2712-9.
31. Borchert D, Federlein M, Ruckbeil O, et al. Prospective evaluation of transvaginal assisted cholecystectomy. *Surg Endosc* 2012;26:3597-604.
32. Brown KM, Moore BT, Sorensen GB, et al. Patient-reported outcomes after single-incision versus traditional laparoscopic cholecystectomy: a randomized prospective trial. *Surg Endosc* 2013;27:3108-15.
33. Bucher P, Pugin F, Buchs NC, et al. Randomized clinical trial of laparoendoscopic single-site versus conventional laparoscopic cholecystectomy. *Br J Surg* 2011;98:1695-702.
34. Burney RE, Jones KR. Ambulatory and admitted laparoscopic cholecystectomy patients have comparable outcomes but different functional health status. *Surg Endosc* 2002;16:921-6.
35. Chen L, Tao SF, Xu Y, et al. Patients' quality of life after laparoscopic or open cholecystectomy. *J Zhejiang Univ Sci B* 2005;6:678-81.
36. Cleary R, Venables CW, Watson J, et al. Comparison of short term outcomes of open and laparoscopic cholecystectomy. *Qual Health Care* 1995;4:13-7.
37. Finan KR, Leeth RR, Whitley BM, et al. Improvement in gastrointestinal symptoms and quality of life after cholecystectomy. *Am J Surg* 2006;192:196-202.
38. Harju J, Paakkonen M, Eskelinen M. Comparison of the quality of life after minilaparotomy cholecystectomy versus laparoscopic cholecystectomy: a prospective randomized study. *Isr Med Assoc J* 2007;9:147-8.
39. Hauters P, Auvray S, Cardin JL, et al. Comparison between single-incision and conventional laparoscopic cholecystectomy: a prospective trial of the Club Coelio. *Surg Endosc* 2013;27:1689-94.
40. Howie MT, Sandblom G, Osterberg J. The impact of pain frequency, pain localization and perceived cause of pain on quality of life after cholecystectomy. *Scand J Gastroenterol* 2017;52:1391-7.
41. Hsueh LN, Shi HY, Wang TF, et al. Health-related quality of life in patients undergoing cholecystectomy. *Kaohsiung J Med Sci* 2011;27:280-8.
42. Johansson M, Thune A, Nelvin L, et al. Randomized clinical trial of day-care versus overnight-stay laparoscopic cholecystectomy. *Br J Surg* 2006;93:40-5.
43. Keus F, de Vries J, Gooszen HG, et al. Laparoscopic versus small-incision cholecystectomy: health status in a blind randomised trial. *Surg Endosc* 2008;22:1649-59.
44. Keulemans Y, Eshuis J, de Haes H, et al. Laparoscopic cholecystectomy: day-care versus clinical observation. *Ann Surg* 1998;228:734-40.
45. Kudsi OY, Castellanos A, Kaza S, et al. Cosmesis, patient satisfaction, and quality of life after da Vinci Single-Site cholecystectomy and multiport laparoscopic cholecystectomy: short-term results from a prospective, multicenter, randomized, controlled trial. *Surg Endosc* 2017;31:3242-50.
46. Kum CK, Eypasch E, Lefering R, et al. Laparoscopic cholecystectomy for acute cholecystitis: is it really safe? *World J Surg* 1996;20:43-8; discussion 48-9.
47. Lamberts MP, Den Ouden BL, Gerritsen JJ, et al. Prospective multicentre cohort study of patient-reported outcomes after cholecystectomy for uncomplicated symptomatic cholelithiasis. *Br J Surg* 2015;102:1402-9.
48. Leung D, Yetasook AK, Carbray J, et al. Single-incision surgery has higher cost with equivalent pain and quality-of-life scores compared with multiple-incision laparoscopic

- cholecystectomy: a prospective randomized blinded comparison. *J Am Coll Surg* 2012;215:702-8.
49. Lirici MM, Califano AD, Angelini P, et al. Laparo-endoscopic single site cholecystectomy versus standard laparoscopic cholecystectomy: results of a pilot randomized trial. *Am J Surg* 2011;202:45-52.
 50. Liu E, Li Z, Wang N, et al. A prospective, randomized, controlled trial of three-port laparoscopic cholecystectomy versus conventional four-port laparoscopic cholecystectomy: is the fourth port really required? *Int J Clin Exp Med* 2016;9:3055-61.
 51. Ma J, Cassera MA, Spaun GO, et al. Randomized controlled trial comparing single-port laparoscopic cholecystectomy and four-port laparoscopic cholecystectomy. *Ann Surg* 2011;254:22-7.
 52. Matovic E, Hasukic S, Ljuca F, et al. Quality of life in patients after laparoscopic and open cholecystectomy. *Med Arh* 2012;66:97-100.
 53. McLean KA, Sheng Z, O'Neill S, et al. The influence of clinical and patient-reported outcomes on post-surgery satisfaction in cholecystectomy patients. *World J Surg* 2017;41:1752-61.
 54. Menten BB, Akin M, Irkorucu O, et al. Gastrointestinal quality of life in patients with symptomatic or asymptomatic cholelithiasis before and after laparoscopic cholecystectomy. *Surg Endosc* 2001;15:1267-72.
 55. Nilsson E, Ros A, Rahmqvist M, et al. Cholecystectomy: costs and health-related quality of life: a comparison of two techniques. *Int J Qual Health Care* 2004;16:473-82.
 56. Pålsson SH, Rasmussen I, Lundstrom P, et al. Registration of health-related quality of life in a cohort of patients undergoing cholecystectomy. *ISRN Gastroenterol* 2011;2011:507389.
 57. Phillips MS, Marks JM, Roberts K, et al. Intermediate results of a prospective randomized controlled trial of traditional four-port laparoscopic cholecystectomy versus single-incision laparoscopic cholecystectomy. *Surg Endosc* 2012;26:1296-303.
 58. Plaisier PW, van der Hul RL, Nijs HGT, et al. Quality of life and the course of biliary and gastrointestinal symptoms after laparoscopic and conventional cholecystectomy. *Dig Surg* 1995;12:87-91.
 59. Planells Roig M, Cervera Delgado M, Garcia Espinosa R, et al. Evaluation of the gastrointestinal quality of life index as a system to prioritize patients on the waiting list for laparoscopic cholecystectomy. *Cir Esp* 2013;91:308-15.
 60. Reibetanz J, Ickrath P, Hain J, et al. Single-port laparoscopic cholecystectomy versus standard multiport laparoscopic cholecystectomy: a case-control study comparing the long-term quality of life and body image. *Surg Today* 2013;43:1025-30.
 61. Rosenmüller MH, Nilsson E, Lindberg F, et al. Costs and quality of life of small-incision open cholecystectomy and laparoscopic cholecystectomy - an expertise-based randomised controlled trial. *BMC Gastroenterol* 2017;17:48.
 62. Rydbeck D, Anesten B, Barje T, et al. Health-Related Quality-of-Life in a cohort undergoing cholecystectomy. *Ann Med Surg (Lond)* 2015;4:22-5.
 63. Saad S, Strassel V, Sauerland S. Randomized clinical trial of single-port, minilaparoscopic and conventional laparoscopic cholecystectomy. *Br J Surg* 2013;100:339-49.
 64. Sadati L, Pazouki A, Tamannaie Z, et al. Quality of life after surgery in candidates of laparoscopic and open cholecystectomy: a comparison study. *Iran Red Crescent Med J* 2016:1-8.
 65. Shi HY, Lee HH, Tsai MH, et al. Long-term outcomes of laparoscopic cholecystectomy: a prospective piecewise linear regression analysis. *Surg Endosc* 2011;25:2132-40.
 66. Squirrell DM, Majeed AW, Troy G, et al. A randomized, prospective, blinded comparison of postoperative pain, metabolic response, and perceived health after laparoscopic and small incision cholecystectomy. *Surgery* 1998;123:485-95.
 67. Sulu B, Yildiz BD, Ilingi ED, et al. Single port vs. four port cholecystectomy-randomized trial on quality of life. *Adv Clin Exp Med* 2015;24:469-73.
 68. Tani M, Kawai M, Okada K, et al. Evaluation of the health-related quality of life for patients following laparoscopic cholecystectomy. *Surg Today* 2015;45:564-8.
 69. Teubner O, Heidecke CD, Kohlmann T, et al. A prospective study comparing quality of life and cosmetic results between single-port and conventional laparoscopic cholecystectomy. *Sci Res* 2016;7:114-25.
 70. Topcu O, Karakayali F, Kuzu MA, et al. Comparison of long-term quality of life after laparoscopic and open cholecystectomy. *Surg Endosc* 2003;17:291-5.
 71. Vetrhus M, Soreide O, Eide GE, et al. Pain and quality of life in patients with symptomatic, non-complicated gallbladder stones: results of a randomized controlled trial. *Scand J Gastroenterol* 2004;39:270-6.
 72. Wagner MJ, Kern H, Hapfelmeier A, et al. Single-port cholecystectomy versus multi-port cholecystectomy: a prospective cohort study with 222 patients. *World J Surg*

- 2013;37:991-8.
73. Wanjura V, Sandblom G. How do quality-of-life and gastrointestinal symptoms differ between post-cholecystectomy patients and the background population? *World J Surg* 2016;40:81-8.
 74. Zapf M, Denham W, Barrera E, et al. Patient-centered outcomes after laparoscopic cholecystectomy. *Surg Endosc* 2013;27:4491-8.
 75. Patrick DL, Guyatt GH, Acquadro C. Chapter 17: Patient-reported outcomes. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011.
 76. 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.
 77. Mueck KM, Cherla DV, Taylor A, et al. Randomized controlled trials evaluating patient-reported outcomes of studies after cholecystectomy: a systematic review. *J Am Coll Surg* 2018;226:183-93.e5.
 78. Cella D, Hahn E, Jensen S, et al. *Patient-Reported Outcomes in Performance Measurement*. Research Triangle Park (NC): RTI Press; 2015.
 79. Enochsson L, Thulin A, Osterberg J, et al. The Swedish Registry of Gallstone Surgery and Endoscopic Retrograde Cholangiopancreatography (GallRiks): A nationwide registry for quality assurance of gallstone surgery. *JAMA Surg* 2013;148:471-8.
 80. Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. *JAMA* 2018;319:483-94.
 81. Streiner DL, Norman GR. *Health measurement scales. A practical guide to their development and use*. 4th edition. New York: Oxford University Press, 2008.

Cite this article as: Daliya P, Gemmill EH, Lobo DN, Parsons SL. A systematic review of patient reported outcome measures (PROMs) and quality of life reporting in patients undergoing laparoscopic cholecystectomy. *HepatoBiliary Surg Nutr* 2019;8(3):228-245. doi: 10.21037/hbsn.2019.03.16

Table S1 Study characteristics (detailed information of each of the 51 non-validation studies)

| Author, year | Study type | No. of centres | N | Diagnosis | Intervention | Control | Study outcome/s |
|---------------------------|----------------------|----------------|-------|---|----------------------------|-----------------------------|---|
| Abd Ellatif, 2013 (24) | RcT | 1 | 250 | Cholelithiasis | SALC | CLC | HRQoL, operative time, pain, cosmesis, blood loss, resumption diet, LOS, DRNA |
| Ainslie, 2003 (25) | RCT | 1 | 40 | Cholelithiasis | MPLC | CLC | Operative time, analgesia, pain, pulmonary function, HRQoL, serum stress response |
| Aspinen, 2017 (26) | RcT | 2 | 109 | Cholelithiasis | MC | CLC | HRQoL |
| Barkun, 1992 (27) | RCT | 5 | 70 | Cholelithiasis | MC | CLC | LOS, DRNA, pain, HRQoL, operative time, conversion, resumption diet |
| Barthelsson, 2008 (28) | RcT | 1 | 73 | Cholelithiasis | Outpatient CLC | Inpatient CLC | Pain, anxiety/distress symptoms, HRQoL |
| Bingener, 2015 (29) | RcT | 1 | 110 | Cholelithiasis | SPLC | CLC | Pain, HRQoL, serum cytokines, heart rate variability |
| Bitzer, 2008 (30) | Longitudinal cohort | 2 | 205 | Cholelithiasis, cholecystitis | CLC | none | HRQoL, symptoms, satisfaction |
| Borchert, 2012 (31) | Observational cohort | 1 | 275 | Cholelithiasis | TVC | CLC | Pain, analgesia, HRQoL, sexual function |
| Brown, 2013 (32) | Randomised cohort | 1 | 79 | Cholelithiasis, biliary dyskinesia | SILC | CLC | HRQoL, LOS, operative costs |
| Bucher, 2011 (33) | RcT | 1 | 150 | Cholelithiasis, cholecystitis, pancreatitis, incidental cholangiocarcinoma | LESS | CLC | Cosmesis, pain, analgesia, morbidity, operative time, port enlargement, HRQoL, LOS, DRNA, operative costs |
| Burney, 2002 (34) | Observational cohort | 1 | 140 | Cholelithiasis | Outpatient CLC | Inpatient CLC | HRQoL |
| Chen, 2005 (35) | Prospective cohort | 1 | 51 | Cholelithiasis, chronic cholecystitis | CLC | OC | HRQoL |
| Cleary, 1995 (36) | Observational cohort | 1 | 391 | Cholelithiasis, cholecystitis | CLC | OC | HRQoL, morbidity |
| Finan, 2006 (37) | Observational cohort | 1 | 104 | Cholelithiasis | CLC | none | HRQoL, symptoms |
| Harju, 2007 (38) | RcT | 1 | 157 | Cholelithiasis | MLC | CLC | HRQoL |
| Hauters, 2013 (39) | Observational cohort | 9 | 104 | Cholelithiasis | SILC | CLC | Operative time, morbidity, HRQoL, cosmesis, satisfaction |
| Howie, 2017 (40) | Observational cohort | 14 | 4,021 | Cholelithiasis | CLC, OC, MLC | none | Pain, symptoms, HRQoL, long-term outcomes |
| Hsueh, 2011 (41) | Observational cohort | 2 | 297 | Cholelithiasis, cholecystitis | CLC, OC | none | HRQoL |
| Johansson, 2006 (42) | RcT | 1 | 107 | Cholelithiasis | Outpatient CLC | Inpatient CLC | HRQoL |
| Keus, 2008 (43) | RcT | 1 | 257 | Cholelithiasis | CLC | SIC | HRQoL, cosmesis |
| Keulemans, 1998 (44) | RCT | 1 | 80 | Cholelithiasis | Outpatient CLC | Inpatient CLC | HRQoL, readmission |
| Kudsi, 2017 (45) | RCT | 8 | 138 | Cholelithiasis, cholecystitis, biliary polyps, biliary dyskinesia | RSSC | CLC | HRQoL, satisfaction, cosmesis, perioperative outcomes |
| Kum, 1996 (46) | Observational cohort | 1 | 478 | Cholelithiasis, cholecystitis | CLC acute cholecystitis | CLC cholelithiasis | Pain, fatigue, HRQoL, morbidity |
| Lamberts, 2015 (47) | Observational cohort | 3 | 552 | Cholelithiasis | CLC | None | Pain, HRQoL, symptoms |
| Leung, 2012 (48) | RcT | 3 | 79 | Choledocholithiasis, cholecystitis, pancreatitis | SILC | CLC | HRQoL, cosmesis, pain, morbidity, cost, efficiency |
| Lirici, 2011 (49) | Observational cohort | 2 | 40 | Cholelithiasis | LESS | CLC | HRQoL, cosmesis, pain, LOS, operative time, conversion, operative difficulty, morbidity |
| Liu, 2016 (50) | RCT | 1 | 245 | Cholelithiasis | 3-port LC | CLC | HRQoL, cosmesis, pain, LOS, costs |
| Ma, 2011 (51) | RCT | 1 | 43 | Cholelithiasis, biliary polyps, biliary dyskinesia | SPLC | CLC | Pain, HRQoL, cosmesis, operative time, LOS, morbidity |
| Matovic, 2012 (52) | Observational cohort | 1 | 120 | Cholelithiasis | CLC | OC | HRQoL |
| McLean, 2017 (53) | Observational cohort | 1 | 234 | Cholelithiasis | CLC, OC | none | HRQoL, satisfaction, symptoms |
| Mentes, 2001 (54) | Observational cohort | 1 | 67 | Cholelithiasis | CLC symptomatic gallstones | CLC asymptomatic gallstones | HRQoL |
| Nilsson, 2004 (55) | RCT | 5 | 726 | Cholelithiasis | MC | CLC | HRQoL, costs |
| Pålsson, 2011 (56) | Observational cohort | 5 | 330 | Cholelithiasis, cholecystitis, pancreatitis, choledocholithiasis, biliary polyps, biliary tumours | CLC | None | HRQoL |
| Phillips, 2012 (57) | RCT | 10 | 197 | Cholelithiasis, biliary polyps, biliary dyskinesia | SILC | CLC | Pain, cosmesis, HRQoL, operative time |
| Plaisier, 1995 (58) | Observational cohort | 1 | 31 | Cholelithiasis | CLC | OC | HRQoL, symptoms |
| Plannells Roig, 2013 (59) | Observational cohort | 1 | 99 | Cholelithiasis, cholecystitis, pancreatitis, choledocholithiasis | CLC, OC, LESS | None | Disease impact, HRQoL |
| Reibetanz, 2013 (60) | Observational cohort | 1 | 100 | Cholelithiasis, cholecystitis | SPLC | CLC | HRQoL, cosmesis, satisfaction, operative time, morbidity, LOS |
| Rosenmüller, 2017 (61) | RCT | 2 | 355 | Cholelithiasis, cholecystitis, pancreatitis, choledocholithiasis | SIC | CLC | HRQoL, costs, pain, morbidity, operative time, LOS, conversion, readmissions |
| Rydbeck, 2015 (62) | Observational cohort | 6 | 919 | Cholelithiasis, cholecystitis, pancreatitis, choledocholithiasis | CLC, OC | None | HRQoL |
| Saad, 2013 (63) | RcT | 1 | 105 | Cholelithiasis | SPLC, mini-laparoscopic LC | CLC | Pain, HRQoL, cosmesis, satisfaction |
| Sadati, 2016 (64) | Observational cohort | 2 | 100 | Cholelithiasis, cholecystitis | CLC | OC | HRQoL |
| Shi, 2011 (65) | Prospective cohort | 2 | 353 | Cholelithiasis | CLC | None | HRQoL |
| Squirrel, 1998 (66) | RcT | 1 | 100 | Cholelithiasis | CLC | SIC | HRQoL, pain, metabolic & respiratory response |
| Sulu, 2015 (67) | RcT | 1 | 60 | Cholelithiasis | SPLC | CLC | HRQoL, satisfaction |
| Tani, 2015 (68) | Observational cohort | 1 | 127 | Cholelithiasis, biliary polyps | CLC | None | HRQoL, pain |
| Teubner, 2016 (69) | Prospective cohort | 2 | 66 | Cholelithiasis | SILC | CLC | HRQoL, cosmesis, satisfaction |
| Topcu, 2003 (70) | Case control | 1 | 200 | Cholelithiasis, biliary polyps, cholecystitis, pancreatitis, choledocholithiasis | CLC | OC | HRQoL, clinical outcomes |
| Vetthus, 2004 (71) | RCT + crossover | 3 | 137 | Cholelithiasis | CLC, OC | Observation | HRQoL, pain |
| Wagner, 2013 (72) | Observational cohort | 1 | 222 | Cholelithiasis, biliary polyps, cholecystitis | SPLC | CLC | HRQoL, morbidity |
| Wanjura, 2016 (73) | Observational cohort | 1 | 451 | Cholelithiasis, biliary polyps, cholecystitis, pancreatitis, choledocholithiasis | Cholecystectomy | None | HRQoL |
| Zapf, 2013 (74) | Observational cohort | 3 | 100 | Cholelithiasis, biliary polyps, pancreatitis, biliary dyskinesia, choledocholithiasis | CLC | None | HRQoL, pain |

RCT, randomised controlled trial; RcT, randomised clinical trial; CLC, conventional laparoscopic cholecystectomy; LC, laparoscopic cholecystectomy; LESS, laparoendoscopic single site cholecystectomy; MC, minilaparotomy cholecystectomy; MPLC, micro-puncture laparoscopic cholecystectomy; OC, open cholecystectomy; RSSC, robotic single-site cholecystectomy; SALC, single-access laparoscopic cholecystectomy; SIC, small incision cholecystectomy; SILC, single incision laparoscopic cholecystectomy; SPLC, single port laparoscopic cholecystectomy; TVC, transvaginal cholecystectomy; LOS, length of stay; DRNA, duration to return to normal activity; HRQoL, health related quality of life.

Table S2 COSMIN (Consensus-based Standards for the selection of health Measurement Instruments) analysis (summary of methodological qualities assessed for each study)

| Study | Eypash, 1995 (18) | Lien, 2007 (19) | Quintana, 2001 (20) | Sandblom, 2009 (21) | Chen, 2006 (22) | Russell, 1996 (23) |
|---|-------------------|-----------------|---------------------|---------------------|-----------------|--------------------|
| PRO instrument | GIQLI | GIQLI | GIQLI | GIQLI | CSQ | GIC |
| A. Internal consistency: Cronbach's alpha (calculated for each subscale). Extent to which items in a (sub) scale are correlated (measuring the same construct) | Y | Y | Y | Y | Y | Y |
| B. Reliability: ICC. Extent to which patients can be distinguished from each other despite measurement errors (proportion of the total variance in measurement properties which is due to 'true' differences between patients) | Y | Y | Y | Y | Y | Y |
| C. Measurement error: systematic and random error of a patient's score that is not attributed to true changes in patient's disease status | x | x | x | x | x | x |
| D. Content validity (face validity & construct validity): degree to which content of instrument is adequate reflection of construct measured (involves patients/experts in development of PROMs through interview, testing, focus groups) | Y | x | x | x | Y | Y |
| E. Structural validity: degree to which scores of instrument are reflection of dimensionality of construct measured (determined by factor analysis) | x | x | x | x | x | x |
| F. Hypotheses testing: degree to which scores of instrument are consistent with hypotheses based on assumption the instrument validly measures what it is meant to (hypotheses must give direction & magnitude) | x | x | x | x | x | x |
| G. Cross-cultural validity: degree to which performance of items on translated questionnaire is an adequate reflection of the performance of items on original questionnaire | x | Y | Y | Y | x | x |
| H. Criterion validity: extent to which scores on questionnaire relate to a gold standard | x | x | x | x | x | x |
| I. Responsiveness: ability of a questionnaire to detect clinically important changes over time (ability to detect improvement/deterioration) | Y | Y | Y | Y | x | Y |
| J. Interpretability: degree to which you can assign qualitative meaning to quantitative scores | x | x | x | x | x | x |

PROs, patient reported outcomes; PROMs, patient reported outcome measures; GIQLI, gastrointestinal quality of life index; CSQ, Otago Gallstones Condition Specific Questionnaire; GIC, Gallstone Impact Checklist; ICC, intraclass correlation coefficient; Y, yes; x, no.