The Newcastle ENDOPREM™: a validated patient reported experience measure for gastrointestinal endoscopy

Laura J Neilson, Linda Sharp, Joanne M Patterson, Christian von Wagner, Paul Hewitson, Lesley M McGregor, Colin J Rees

ABSTRACT

Objectives Measuring patient experience of gastrointestinal (GI) procedures is a key component of evaluation of quality of care. Current measures of patient experience within GI endoscopy are largely clinician derived and measured; however, these do not fully represent the experiences of patients themselves. It is important to measure the entirety of experience and not just experience directly during the procedure. We aimed to develop a patient-reported experience measure (PREM) for GI procedures.

Design Phase 1: semi-structured interviews were conducted in patients who had undergone GI endoscopy or CT colonography (CTC) (included as a comparator). Thematic analysis identified the aspects of experience important to patients. Phase 2: a question bank was developed from phase 1 findings, and iteratively refined through rounds of cognitive interviews with patients who had undergone GI procedures, resulting in a pilot PREM. Phase 3: patients who had attended for GI endoscopy or CTC were invited to complete the PREM. Psychometric properties were investigated. Phase 4 involved item reduction and refinement.

Results Phase 1: interviews with 35 patients identified six overarching themes: anxiety, expectations, information & communication, embarrassment & dignity, choice & control and comfort. Phase 2: cognitive interviews refined questionnaire items and response options. Phase 3: the PREM was distributed to 1650 patients with 799 completing (48%). Psychometric properties were found to be robust. Phase 4: final questionnaire refined including 54 questions assessing patient experience across five temporal procedural stages.

Conclusion This manuscript gives an overview of the development and validation of the Newcastle ENDOPREM™, which assesses all aspects of the GI procedure experience from the patient perspective. It may be used to measure patient experience in clinical care and, in research, to compare patients’ experiences of different endoscopic interventions.

INTRODUCTION

Around one-third of the population will undergo gastrointestinal (GI) endoscopy during their lifetime, with approximately 1.5 million endoscopies performed annually in England and 17 million in the USA. CT colonography (CTC) is performed in many countries as an alternative to colonoscopy or where colonoscopy is incomplete. There is significant variation in CTC use between countries with around 150 000 CTCs undertaken in England annually, and although less...
widely used in the USA, it is performed in at least 700 centres.\(^\text{15}\)

Consistent high-quality endoscopy is essential to ensure that procedures are safe and that lesions are not missed, however, there is significant and unacceptable variation in practice and quality.\(^\text{6,10}\) Patient experience is one of the key dimensions of high-quality clinical care. Increasing emphasis on patient experience as a marker of quality has emerged in recent years because of the recognition that positive patient experience is associated with better patient outcomes.\(^\text{11}\) In England, a report into a failing hospital found that losing sight of patients at the centre of healthcare was a key component of system failures.\(^\text{12}\) As a result, patient experience is now routinely assessed across many aspects of primary and secondary care services in the English National Health Service. Similarly, in the USA, there is increasing use of experience measures (such as the Hospital Consumer Assessment of Healthcare Providers and Systems Survey) to evaluate healthcare providers.\(^\text{13}\) The value of such experience measures (such as the Hospital Consumer Assessment of Healthcare Providers and Systems Survey) to evaluate healthcare providers.\(^\text{13}\) The value of such information is that it lets healthcare providers identify which aspects of a service may need attention or require improvement.

Within the field of endoscopy, there is evidence that colonoscopists who deliver high-quality procedures also deliver better patient experience.\(^\text{14}\) Patient experience is a key determinant of participation in bowel cancer screening programmes and influences whether patients attend for repeat disease surveillance procedures.\(^\text{11,15}\)

Current measures of patient experience within the field of GI endoscopy focus almost completely on procedural aspects such as pain or discomfort but ignore elements such as preprocedural communication, anxiety and preparation for procedures.\(^\text{16}\) Additionally, almost all measures are clinician derived and measured. For example, the Gastrointestinal Endoscopy Satisfaction Questionnaire focuses on satisfaction and did consult with patients during question development, however, these questions were developed based on literature review and were not generated by the patients themselves following patient interviews using best practice for development of measures of patient experience.\(^\text{17}\) It has been reported that patients and clinicians prioritise different aspects of experience with clinicians prioritising the procedure itself, including aspects such as comfort, whereas communication and interaction with the endoscopist are more important to patients.\(^\text{18}\) Elements of the endoscopy experience other than those at the time of the procedure, for example, bowel preparation and preprocedural anxiety, are also important. Preprocedural anxiety has been found to be a key determinant of pain and discomfort experience during other clinical procedures.\(^\text{19}\) Moreover, measuring only procedural comfort scores gives information on only a very small part of the overall patient experience. Patient experience of CTC is not routinely measured in clinical practice and its measurement in research settings tends to focus on comfort and bowel preparation.\(^\text{20,21}\)

Endoscopy organisations advise measuring patient experience but provide little advice on how this should be done. The European Society of Gastrointestinal Endoscopy (ESGE) and British Society of Gastroenterology (BSG) have highlighted the importance of measuring and improving patient experience and recommend measuring self-reported patient experience using a ‘validated scale,’ however, they acknowledge the lack of a standardised approach.\(^\text{22,25}\) Patient-reported experience measures (PREMs) are tools to measure patients’ experience of healthcare and should address those aspects of care defined by patients as being important to them; thus, patients should be involved in developing such measures and in defining what they measure.\(^\text{24,25}\) The systematic development and validation of PREMs have been undertaken in a range of specialties and disease conditions, for example, cancer care, paediatric emergency care and sickle cell disease but not in the field of GI endoscopy or CTC.\(^\text{26–28}\) The approach to devising a PREM includes literature review to identify current tools and aspects contributing to patient experience; a qualitative stage involving relevant patients to identify areas of care important to them; questionnaire design informed by the qualitative work; cognitive testing of the questionnaire and field testing to assess validity.\(^\text{25,26}\) Tools to measure the total, or overall, experience of GI endoscopy do not exist. Those tools which have been developed to measure some elements of experience have not been systematically developed with patients. This paper describes the rigorous process of developing the first patient-derived, validated, PREM for GI endoscopy.

METHODS
The study was undertaken in the North-East of England. The development of the PREM involved four phases and was undertaken according to the established COSMIN (COnsensus-based Standards for the selection of health status Measurement INstruments) criteria for development of health measurement tools.\(^\text{30}\)

Phase 1: concept elicitation
This phase provided in-depth and detailed identification and description of the experiences of patients who had undergone oesophagogastrooduodenoscopy (OGD), colonoscopy or CTC and has been reported in detail elsewhere.\(^\text{31}\) Between February 2016 and April 2017, patients aged ≥18 years who had undergone one of these procedures for symptoms or surveillance (but not within the national bowel cancer screening programme) in one hospital were invited to participate in one-to-one semi-structured interviews. Eligible patients on a series of endoscopy lists were approached in the department before their procedure. Purposive sampling was used to ensure a range of procedures, age and sex among participants. These were guided by a topic guide developed from literature review and expert opinion; the guide was used flexibly so that any new issues identified
by patients were added to the guide to be explored in subsequent interviews. Interviews were audio-recorded and transcribed. Recruitment continued until data saturation. Thematic analysis was undertaken to identify overarching themes. This phase identified aspects of experience that mattered to patients to inform questionnaire content. It was also used to identify whether it was possible to develop a PREM that would apply across different GI procedures.

Phase 2: development and content validity
This phase was undertaken between April and August 2017 and involved iterative development of the pilot PREM. Informed by the qualitative interviews and a focused literature review, a question bank was generated. This involved decisions on topics to be included, their order, question format and response format. Several rounds of revision and review were undertaken by the research team.

The draft questionnaire—designed for patient self-completion—was then developed and pretested, using several rounds of face-to-face cognitive interviews with patients who had undergone GI procedures, from June to July 2017. Interviewees were invited to ‘talk aloud’ as they completed the questionnaire, with verbal probing used to clarify any problems or issues. Different styles of question format and layout were tested. Interviews were audio-recorded and transcribed. Analysis after each round used a systematic approach to identify problems. The questionnaire was refined after each round and the refined draft tested in the subsequent round. Once no new issues arose in interviews, the pilot PREM was agreed. This process ensured patient comprehension of the questions and face and content validity of the questionnaire (ie, confirmed relevance and comprehensiveness of the questions).

Phase 3: psychometric properties and validation
This phase evaluated the psychometric properties of the pilot questionnaire and validated it in a large cohort. Between October 2017 and September 2018, the pilot PREM was given to patients following GI endoscopy or CTC in four English NHS hospital Trusts. The sample size was based on the number of questions included in the pilot PREM and exceeded the recommended 15 completed questionnaires per item for studies of psychometric qualities. Patients were asked to take it home and complete and return it in the prepaid envelope provided. A patient information sheet (PIS) was included with the questionnaire pack and it was made clear to participants that return of a completed questionnaire was deemed to signify consent.

Statistical analysis focused on response rates and patterns, missing values, ‘floor’ and ‘ceiling’ effects (ie, propensity of respondents to endorse the extreme ends of the response scale) and correlations between question responses (to identify agreement between items and question redundancy) and with the total score (sum of all the experience questions). Analysis was performed using IBM® SPSS® V.24 (IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp).

Following completion of Phase 3, all questions were reviewed systematically alongside the Phase 3 data to determine where refinements in wording or layout were required. Those questions considered to be redundant were removed, for example, those which correlated poorly with others. This resulted in the Newcastle ENDOPREM™

RESULTS
Phase 1: concept elicitation
Thirty-five patients (15 OGD, 10 colonoscopy, 10 CTC) were interviewed. Characteristics of these patients are

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of participants (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>19 (54.3)</td>
</tr>
<tr>
<td>Female</td>
<td>16 (45.7)</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
</tr>
<tr>
<td>≤54</td>
<td>12 (34.3)</td>
</tr>
<tr>
<td>55–64</td>
<td>6 (17.1)</td>
</tr>
<tr>
<td>65–74</td>
<td>12 (34.3)</td>
</tr>
<tr>
<td>≥75</td>
<td>5 (16.3)</td>
</tr>
<tr>
<td><strong>Procedure</strong></td>
<td></td>
</tr>
<tr>
<td>OGD (including transnasal endoscopy)</td>
<td>15 (42.8)</td>
</tr>
<tr>
<td>Colonoscopy</td>
<td>10 (28.6)</td>
</tr>
<tr>
<td>CTC</td>
<td>10 (28.6)</td>
</tr>
<tr>
<td><strong>Indication</strong></td>
<td></td>
</tr>
<tr>
<td>Alarm symptoms</td>
<td>12 (34.3)</td>
</tr>
<tr>
<td>Routine symptoms</td>
<td>17 (48.6)</td>
</tr>
<tr>
<td>Surveillance procedures</td>
<td>6 (17.1)</td>
</tr>
</tbody>
</table>
listed in Table 1. When invited to discuss their experience, a few participants focused on a single aspect of the experience which was most important to them: for example, the taste and volume of the bowel preparation or the effect of good communication from staff. However, most participants described their experience in chronological order, that is, as stages of a process, starting with the referral process (henceforth, called ‘before attending for the test’), then visiting the hospital (‘at the hospital, before the test’), undergoing the procedure itself (‘during the test’) and what happened afterwards (‘after the test’). Six themes emerged: anxiety, expectations, information & communication, embarrassment & dignity, comfort and choice & control. All six themes were demonstrated in more than one procedural stage. All themes and procedural stages emerged in relation to all three GI procedures.

The themes were organised by procedural stage, as described in phase 1. Questions were developed for each relevant theme and stage; where possible, language echoed that used by patients in phase 1. Most questions were in the form of statements (eg, ‘during the test, my dignity was maintained at all times’) with a five-level Likert-type response scale, ranging from ‘strongly agree’ to ‘strongly disagree’. To minimise response set bias (where respondents endorse the same response option for all questions), both positive and negative questions were included (ie, some statements were phrased such that ‘strongly agree’ would indicate a ‘positive’ experience, and others phrased so that ‘strongly disagree’ would indicate a ‘positive’ experience).

Ten rounds of review and revision were conducted by the study team. Five rounds of cognitive testing were undertaken, with three endoscopy patients in each round (15 participants in total, Table 2). Focused literature review found no additional aspects of patient experience not covered by the pilot PREM.

The pilot PREM thus comprised 59 questions which spanned patient experience; plus a series of questions on respondent characteristics (eg, age, ethnicity) as these have been found to be associated with patient experience in many clinical contexts; plus three questions on potential ‘explanatory’ factors (which might affect patient experience; eg, endoscopist gender); and a further two questions on ‘overall’ experience. The sections of the pilot PREM were section A—completing this survey (10 questions), section B—before coming to hospital for your test (16 questions), section C—preparing for your test (six questions), section D—at the hospital, before the test (five questions), section E—during the test (21 questions), section F—after the test (11 questions) and section G—overall experience (three questions).

### Phase 3: psychometric properties and validation

The PREM underwent multisite validation and was given to 1650 eligible patients of whom 799 responded (response rate=48.4%). The response rate was higher in older patients and those undergoing lower GI procedures were more likely to respond (Table 3). As patients took the PREM away to complete, the time to complete it was not measured; however, in phase 2, the interview time was 10–15 min. Respondents’ ages ranged from 18 to 95, with a mean age of 65.3 (SD 12.6). 43.3% of respondents were male and the majority (98.4%) were of white British ethnicity. 41.1% of respondents underwent OGD (including 0.8% who underwent transnasal endoscopy), 43.3% underwent colonoscopy, 1.3% underwent both OGD and colonoscopy on the same day (referred to as ‘OGD & colonoscopy’ henceforth) and 14.4% underwent CTC.

There was notable variation in the numbers of participants recruited per procedure type; 10 patients who had both OGD & colonoscopy performed responded to the questionnaire (of 34 invited) and 115 CTC participants...

Open access

(4 of 209 invited). Far fewer CTCs are done at each of the sites than OGD and colonoscopies, and this was reflected in the number of participants recruited in this group.

There was a discrepancy between the procedure the patient reported that they had attended for and the procedure confirmed by the clinical team in 53 cases (6.6%). The section that should have been completed by those undergoing colonoscopy and CTC only (covering issues around bowel preparation) was completed by 96.1% of participants to whom it applied and 18.0% of those to whom it did not.

Completion rates of individual questions were high; for only three questions, more than 5% of people failed to provide an answer. The oldest age group (>75 years) was significantly more likely to miss questions.

In terms of potential question redundancy, two pairs of questions correlated strongly (rs>0.8) and four questions poorly correlated with any others (rs<0.3). Eight questions, including the four which poorly correlated with any others, had poor corrected item-total correlation (ITC <0.3), meaning that they did not correlate well with the overall sum of all questions.

Twenty-five questions had a ceiling effect (>40% of respondents endorsed the ‘best’ response). No questions had floor effects (>40% choosing the ‘worst’ option).

**Phase 4: item reduction and refinement**

Five questions were removed as they were considered redundant because of poor inter-item correlation and ITC. Two items (which were strongly correlated) were merged; they asked a similar question with slightly different time points. One question was added by the study team to simplify an item and one explanatory question was rephrased with a further explanatory question added for clarity.

The final Newcastle ENDOPREM™ includes 10 demographic/patient characteristic questions, 54 patient experience and four explanatory questions (online supplemental file 1). A shorter version will be available for specific procedures, for example, without the bowel preparation section for upper GI endoscopy.

**DISCUSSION**

Understanding and improving patient experience are fundamental to delivering high-quality GI endoscopy. We give an overview of a systematic and rigorous approach to developing the first validated and fully patient-derived PREM for GI endoscopy, the Newcastle ENDOPREM™, conducted according to the COSMIN criteria.

The Newcastle ENDOPREM™ is robust and comprehensive and in this form is designed to capture experience across the entire patient journey, not simply during the procedure itself. In its current form, it is applicable to upper GI endoscopy, colonoscopy and CTC. The various elements of the patient journey—including how the patient received their appointment, how they experienced the facilities in which they waited for the procedure, how results were given and what would happen next were conveyed—were all things that patients raised in the qualitative work as being important to them.31

A range of measures are available to measure individual components of the endoscopy process. For example, a number of procedural comfort scores exist and some of these have been validated.35 However, almost all measures are clinician or expert designed rather than being developed based on what patients report as being important. The Global Rating Scale patient experience domain is used internationally and was derived from literature review and expert opinion.36 The Gastrointestinal

---

**Table 3 Phase 3 response rates according to participant characteristics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Respondents n (%)</th>
<th>Non-respondents n (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>346 (49.7)</td>
<td>350 (50.3)</td>
<td>0.342</td>
</tr>
<tr>
<td>Female</td>
<td>444 (47.3)</td>
<td>494 (52.7)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤54</td>
<td>142 (30.4)</td>
<td>325 (69.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>55–64</td>
<td>191 (48.7)</td>
<td>201 (51.3)</td>
<td></td>
</tr>
<tr>
<td>65–74</td>
<td>271 (60.4)</td>
<td>178 (39.6)</td>
<td></td>
</tr>
<tr>
<td>≥75</td>
<td>181 (56.4)</td>
<td>140 (43.6)</td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OGD</td>
<td>328 (45.6)</td>
<td>392 (54.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Colonoscopy</td>
<td>346 (54.3)</td>
<td>291 (45.7)</td>
<td></td>
</tr>
<tr>
<td>CTC</td>
<td>115 (55.0)</td>
<td>94 (45.0)</td>
<td></td>
</tr>
<tr>
<td>OGD and colonoscopy</td>
<td>10 (29.4)</td>
<td>24 (70.6)</td>
<td></td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>98 (43.9)</td>
<td>125 (56.1)</td>
<td>0.045</td>
</tr>
<tr>
<td>B</td>
<td>193 (53.8)</td>
<td>166 (46.2)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>220 (50.1)</td>
<td>219 (49.9)</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>288 (45.7)</td>
<td>342 (54.3)</td>
<td></td>
</tr>
</tbody>
</table>

P value for χ² tests.

CTC, CT colonography; OGD, oesophagogastroduodenoscopy.
Endoscopy Satisfaction Questionnaire was developed by including the ‘most relevant’ questions from existing measures. Although patients were involved to ensure face and content validity, they were not involved in initially defining the questions. While it covers some preprocedural and postprocedural items, elements about preparation for the test or later results are not covered in detail. The Gastronet is a clinician-derived questionnaire for assessing patient experiences of endoscopy in Europe. It includes three questions about discomfort and a question about satisfaction with the information given about the test and results, which patients complete the day following their procedure. In the USA, the American Society for Gastrointestinal Endoscopy recommends that a general satisfaction scale is used, modified for endoscopic procedures. This is not specific to GI procedures and was developed without patient input, nor does it assess comfort. The Newcastle ENDOPREM™ addressed the assessment of patient experience differently by first establishing the aspects of experience that mattered to endoscopy patients, then systematically developing and testing a tool to measure all of those aspects of experience, while adhering to best practice in health measurement development according to the established COSMIN checklist. Patients were also involved in the process of refining and improving the instrument and in the psychometric validation study. Thus, the instrument is both grounded in patient experience and seeks to measure that experience.

The length of this questionnaire is an important issue to be considered. The PREM is necessarily long as it covers the entirety of the patient experience and covers upper and lower GI procedures with and without sedation. However, despite this length, the completion time was around 10–15 min in the cognitive interview round. In the validation survey, the response rate of 48% is comparable to similar self-completion questionnaires in other areas of clinical practice. Completion rates might be affected by when the questionnaire is administered. We took the decision to ask patients to complete the questionnaire after leaving the department to avoid the possibility that their ratings of experience would be influenced either by sedation or by still being present in the hospital within the department that had provided them with care. There are different possible modes and timing of questionnaire administration (eg, paper vs digital, provided at end of clinic appointment vs sent by post after discharge) and these will be the focus of further work. Were the questionnaire too long then the completion rate of questions might be expected to be impacted, however, the completion of individual questions was excellent with very low levels of missing data. Moreover, there was no evidence that the number of questions missed increased in later parts of the questionnaire. The full questionnaire covers upper and lower GI endoscopy and CTC. At the outset, it was unclear whether it would be possible to develop a PREM that was applicable to patients who had undergone different endoscopic procedures. The work here demonstrates that the issues that matter to patients are very similar across GI procedures. The main difference—as would be expected—is the issue of bowel preparation for colonoscopy, which patients described in rich and varied ways in the qualitative interviews. We designed the instrument so that the questions relating to bowel preparation were in a contained section, however, we note that this section was completed by 18% of those to whom it did not apply. When applied to clinical practice or research, only the relevant parts of the questionnaire will be distributed and, thus, the PREM will be significantly shorter and clearer.

The PREM will also need to be modified where clinical practice varies. In the UK, conscious sedation is used for colonoscopy. In other countries, for example, USA and Australia, deep sedation is used. Therefore, non-relevant sections of the PREM, for example, those covering intraprocedural pain or comfort, will not be completed where they are not needed. While redundant sections of the PREM will not be given to particular patients and in particular settings, it is important that this is based on only removing those irrelevant sections and not simply removing sections that clinicians or researchers consider of lesser importance.

The inclusion of CTC within this PREM was to allow a comparator from a non-endoscopy GI investigation, which may be considered an alternative to colonoscopy. Much of the argument for the role of CTC relates to experience, so developing a PREM to allow this to be measured accurately was one of the goals of this research.

A potential weakness of the study is the study population. Phases 1 and 2 were undertaken in an expert centre with a strong track record in endoscopy and endoscopy research. Phase 3 was undertaken across four sites to increase diversity of patients and experiences, but these were all in the North-East of England, an area with very limited ethnic diversity. Very few participants reported their ethnic group as anything other than White British (and only 0.5% of the participants in phases 1 and 2 were non-White British). This means we cannot be certain that the performance and properties of the questionnaire are the same in non-White British ethnic groups.

Further work is planned to test the PREM further in international settings and settings of wider ethnic diversity. We also currently lack information on the suitability of the questionnaire for patients who had other types of endoscopic procedures (eg, flexible sigmoidoscopy); testing this is also important.

There are two possible uses of the Newcastle ENDOPREM™. As a patient experience measure, the detail and granularity of the questions across the journey can help healthcare providers identify which specific aspects of their service may benefit from improvement. The instrument may also be used to compare patient experiences of different endoscopic interventions in research studies. Detailed understanding of the different components of endoscopy and how these affect patient experience would be valuable in the context of such head-to-head comparisons.
The Newcastle ENDOPREM™ is currently being adapted for other GI procedures, including capsule endoscopy and CytoSponge. These procedures are being introduced into routine clinical practice, in some instances as an alternative to endoscopy, and, therefore, the Newcastle ENDOPREM™ will be available to compare experiences of these procedures, both in the clinical and research setting.

The Newcastle ENDOPREM™ has been developed through a robust and comprehensive pathway. The themes are likely to remain constant despite the changingendoscopic landscape. This PREM was developed prior to the novel COVID-19 pandemic and while delivery of endoscopy may change, and consequently how patients rate their experience, this tool should remain a valid way to measure that experience.

CONCLUSION

The Newcastle ENDOPREM™ is the first patient-derived PREM that can be used to assess experience of patients who have undergone different GI procedures. It is now available for use in GI endoscopy research or evaluation of routine care.

REFERENCES