Journal of the Canadian Association of Gastroenterology, 2020, 3(2), 91–95
doi: 10.1093/jcag/gwy073
Original Article
Advance Access publication 21 December 2018

OXFORD

Original Article

Validation of the St. Paul's Endoscopy Comfort Scale (SPECS) for Colonoscopy

Jennifer Telford, MD, MPH, FRCPC[1], Iran Tavakoli, MD, MSc[2], Oliver Takach, MSc[1], Ricky Kwok, MD[2], Natasha Harris, BSc[1], Jordan Yonge, BSc[1], Cherry Galorpart, MD[1], Scott Whittaker, MD, FRCPC[1], Jack Amar, MD, FRCPC[1], Gregory Rosenfeld, MD, MHSc, FRCPC[1], Hin Hin Ko, MD, FRCPC[1], Eric Lam, MD, MSc, FRCPC[1], Alnoor Ramji, MD, FRCPC[1], Brian Bressler, MD, MS, FRCPC[1], Robert Enns, MD, FRCPC[1]

[1]Department of Internal Medicine, Division of Gastroenterology, University of British Columbia, Vancouver, British Columbia, Canada; [2]Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada

Correspondence: Jennifer J. Telford, MD, MPH, FRCPC, Pacific Gastroenterology Associates 770–1190 Hornby Street Vancouver, British Columbia, V6Z 2KS, E-mail: jtelford@telus.net

Abstract

Aims: Patient comfort during colonoscopy is an important measure of quality, which can improve patient satisfaction and compliance with future procedures. Our aim was to develop and validate a pain assessment tool based on objective behavioural cues tailored to outpatients undergoing colonoscopy: St. Paul's endoscopy comfort score (SPECS).

Methods: A single-centre, prospective study was conducted in consecutive adults undergoing planned outpatient colonoscopy. Patient comfort was independently assessed by the physician, nurse and a research assistant (observer) using the SPECS and the Gloucester scale (GS). In addition, the nurse-assessed patient comfort score (NAPCOMS), nonverbal pain Assessment tool (NPAT) and Richmond agitation sedation scale (RASS) were completed by the observer. Data on subject demographics, sedation dose and duration of the procedure were collected. Following the procedure, patients completed a patient satisfaction questionnaire, including a visual analogue scale (VAS) to measure their overall perceived pain during the procedure.

Results: The study enrolled 350 subjects. The SPECS showed excellent inter-rater reliability among all three raters with an intra-class coefficient (ICC) of 0.81 (95% CI, 0.78–0.84), while the GS showed good reliability with an ICC of 0.77 (95% CI, 0.73–0.80). The SPECS demonstrated moderate agreement with the patient-reported VAS ratings.

Conclusions: The St. Paul's endoscopy comfort score was successfully validated, demonstrating excellent inter-rater reliability.

Keywords: *colonoscopy; inter-rater reliability; patient comfort scale; validation*

Colonoscopy is used for the diagnosis and treatment of colonic lesions and screening and surveillance of colorectal neoplasia. Most Canadian provinces and territories have organized colon screening programs, with colonoscopy recommended for any participant with an abnormal fecal occult blood test. (1). Performance of high-quality colonoscopies is essential. Quality is largely dependent on the expertise and technical skill of the endoscopist and should be quantitatively and reliably measured. Certain quality indicators, such as adenoma detection rate and adequacy of bowel preparation, are accepted, but assessment of

Received: April 25, 2018; Accepted: December 1, 2018

patient satisfaction is not as well established. (2, 3).

Our group has previously shown that comfort during colonoscopy is associated with improved patient satisfaction and compliance with future procedures (4). Two patient comfort scores commonly used for endoscopy include the La Crosse (WI) (5), and the nurse-assessed patient comfort score (NAPCOMS) (6). The La Crosse (WI) intra-endoscopy sedation comfort score showed moderate inter-reporter validity but a poor correlation with patient satisfaction post-procedure (5). The NAPCOMS was developed for outpatients undergoing colonoscopy with minimal or moderate sedation and showed very good intra-class correlation for overall score, intensity, frequency and duration of pain (6). The modified Gloucester scale (GS) (7), although not formally validated, is routinely used to assess patient comfort during colonoscopy and is easily completed in clinical practice. In addition, with both of these scales, assessment of pain is not specified by objective behavioural cues but rather relies on the observers' subjective perception of the patient's pain.

There is a need for a validated, easy-to-use method to assess patients' pain during colonoscopy. While the NAPCOMs has been validated, feedback from our endoscopy nurses indicated that it was cumbersome to use and did not incorporate nonverbal cues they had used in critical care and postoperative recovery. We developed a scoring tool based on published nonverbal patient comfort scores used in critical care (8, 9). The St. Paul's endoscopy comfort score (SPECS) is tailored for colonoscopy with mild or moderate sedation and includes the frequency of verbal cues, body positioning and anxiety levels with descriptions for each variable. The objective of this study was to determine the inter-observer reliability for the SPECS compared with the GS and whether SPECS correlated with patient-reported outcome measures.

METHODS

The SPECS was developed through review of existing patient comfort scores with physicians who perform colonoscopy and registered nurses who assist in colonoscopy. Several iterations were presented, revised after input from physician and nurse stakeholders, and then trialed during colonoscopy before finalization of the scoring tool (Table 1).

The score was validated prospectively in consecutive patients attending St. Paul's hospital for outpatient colonoscopy from June 17, 2014, to August 15, 2014. Adults age 19 years or older who were capable of reading and understanding English were included. Patients undergoing esophagogastroduodenoscopy and colonoscopy in the same appointment were excluded. Patients underwent colonoscopy in the usual fashion. The patient comfort scale was independently completed by the colonoscopist, the nurse and a research assistant (observer). Approximately 30 minutes following colonoscopy, patients were asked to complete a patient satisfaction survey that included a visual analogue scale (VAS) to record the patient's overall perceived pain during the procedure. Patient age, gender, weight, sedation administered, procedure time and the identity of colonoscopist and assisting nurse were also recorded.

Patient Comfort Assessment Tools

Four patient comfort assessment tools were used: the SPECS (Table 1), the modified GS (Table 2), the NAPCOMS (Table 3), and the nonverbal pain assessment tool (NPAT) (Table 4). The patient's sedation level was assessed using the Richmond agitation sedation scale (RASS) (10). The SPECS and GS were independently completed by the colonoscopist, the nurse and a research assistant (observer). The NPAT, NAPCOMS and RASS were completed by the observer only.

The SPECS for colonoscopy includes three categories: vocalization, position/body language, and patient anxiety/emotion. The frequency (from zero to three) of each category is assessed using specific behavioural criteria, with the total score ranging from zero to nine. The GS is a common, although not validated, global rating (from one to five) of the patient's overall comfort during the procedure. The NPAT is a patient comfort assessment validated in verbal and nonverbal critical care patients (9) and includes five categories: emotion, movement, verbal cues,

Table 1. Saint Paul's Endoscopy Comfort Scale (SPECS) for Colonoscopy

Vocalization: signs of whimpering, moaning, grunting, or vocalized pain complaint				
Frequency	None (relaxed)	1–4	5–9	≥10 or screaming/ crying out
Score	0	1	2	3
Positioning/body language – signs of tensing/guarding due to pain/clutching/leg movements				
Frequency	None (relaxed)	1–4	5–9	≥10 or pronounced and agitated movements
Score	0	1	2	3
Patient anxiety/emotion				
Frequency	None (relaxed)	Slightly agitated; minimal anxiety	Visibly upset, ***can*** be calmed by vocal reassurance	Visibly upset or crying, ***cannot*** be calmed by vocal reassurance
Score	0	1	2	3

Total score:_____ + _____ + _____ = _____

facial cues and position/guarding. Each item is scored from zero to two with a total ranging from zero to 10. This scale has not been validated in endoscopy. The NAPCOMS, which has been validated in the colonoscopy setting (6), includes intensity, frequency and duration of pain (from zero to three), with the total score ranging from zero to nine. The RASS assesses the patient's sedation level based on a 10-point scale, ranging from −5 being unrousable to +4 being combative and is tailored for use in critical care facilities.

Patient-Reported Outcome Measures

The postcolonoscopy satisfaction questionnaire was a modified version of the nine-item Group Health Association of America patient satisfaction survey (4) and was accompanied by a 10-cm visual analogue scale (VAS) measuring the patient's report of pain during colonoscopy.

Statistical Analysis

The primary endpoint of this study is the inter-observer reliability of the SPECS and modified Gloucester scale assessed with the intra-class coefficient (ICC). An accepted interpretation of ICC includes the following: values ≤0.4 demonstrate poor correlation, 0.41–0.59 is fair, 0.60–0.74 is good, and ≥0.75 is excellent reliability (11). The inter-observer reliability of the three categories included in SPECS was assessed using the weighted Kappa statistic because the ICC is not appropriate when examining a small range (i.e., three-point scale). A common interpretation for Kappa is divided as follows: <0, no agreement; 0–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and ≥0.81, almost perfect agreement (11).

The sample size required was estimated from the previously published validation of the NAPCOMS (6), plus adding 15% to account for multiple statistical comparisons.

The correlation between the patient comfort scores (SPECS, GS, NPAT and NAPCOMS) and patient self-reported pain on the VAS was assessed using Spearman's rank correlation (ρ). Given that the scales considered have different minimum and maximum scores, in order to assess the agreement between scales, each score was standardized by converting it into a percentage of maximum possible score (ranging from 0% to 100%) (12). Mean difference between VAS and the patient comfort scores were then computed. A positive difference suggested that the patient comfort score was larger than VAS. Spearman's rank correlation was used to determine whether there was a relationship between patient-reported pain on the VAS and the amount of sedation, RASS and length of procedure.

A waiver of prior consent was granted by the St. Paul's Hospital ethics board to ensure patient's behavioural cues during colonoscopy were not influenced by their knowledge of the study.

Table 2. Modified Gloucester Scale

1	No	No discomfort – resting comfortably throughout
2	Minimal	One or two episodes of mild discomfort, well tolerated
3	Mild	More than 2 episodes of discomfort, adequately tolerated
4	Moderate	Significant discomfort experienced several times during the procedure
5	Severe	Extreme discomfort, experienced frequently during the procedure

Adapted from reference (7).

RESULTS

Three hundred fifty subjects were enrolled, with 33 subjects subsequently excluded due to incomplete questionnaire or patients' declining to participate when asked to complete the

Table 3. Nurse-Assessed Patient Comfort Score (NAPCOMS)

Domain	Item	0	1	2	3	Score
Pain	1- Intensity	None or minimal	Mild	Moderate	Severe	
	2 – Frequency	None	Few (1–2 episodes)	Several (3–4 episodes)	Frequent (>4 episodes)	
	3- Duration	None	Short (<30 sec)	Moderate (30 sec – 1 min)	Long (>1 min)	
	Total Pain Score (Intensity + Frequency + Duration)					
Sedation	Level of Consciousness	Alert	Sleepy but initiates conversation	Responds only when asked or stimulated	Unresponsive or only responds with pronounced stimulation	
Global	Tolerability	Very well tolerated	Reasonably well tolerated	Just tolerated	Poorly tolerated	

Table 4. Nonverbal Patient Assessment Tool (NPAT)

	NONVERBAL PATIENT ASSESSMENT SCORE	
YES	Is the patient able to make vocalizations or sound cues?	NO
	Score under the YES or NO category and total scores	
	EMOTION	
	An effective response to a situation	
0	Smiling, calm, relaxed or none due to coma state or analgesia	0
1	Anxious; irritable; withdrawn; closes eyes; does not engage with physical environment	1
2	Tearful/crying or uncooperative	2
	MOVEMENT	
	Change in placement and positioning of the body and extremities when not engaged in any care activities	
0	None; sleeping comfortable; no unusual movements; or none due to coma state or analgesia	0
1	Restless or slow; decreased movement; reluctant to move; muscle tenseness	2
2	Rigidity; increasing motion; stiffening; tossing; turning; flapping of arms;	3
	VERBAL CUES	
	Sound cues or vocalization other than speech	
0	No vocalization	Not Applicable
1	Whispering; moaning; sighing	
2	Screaming; crying out	
	FACIAL CUES	
	Expression on face	
0	Relaxed, calm expression or none due to coma state or analgesia	0
1	Draws around the mouth or eyes; narrowed eyes	1
2	Wincing; grimacing; clenched teeth; furrowed brow; tightened lip	2
	POSITIONING/GUARDING	
	Body responses that imply a protection of the body from contact with external touch	
0	Relaxed body or none due to coma state or analgesia	0
1	Guarding/ tense	2
2	Jumpy when touched; clutching the side-rails; withdraws when touched	3
	TOTAL	

questionnaire after the procedure. The mean age of the 317 subjects was 59.2 years (SD 13.3 years), and 52% of the subjects were female. The average dose of sedation was 3.7 mg (SD 1.3 mg) of midazolam and 64.5 mcg (SD 28.0 mcg) of fentanyl. Nine (2.8%) patients chose not to receive sedation. Mean RASS sedation score was −1.4 (SD 1.0) with 89% of subjects lightly sedated during the colonoscopy (RASS score of 0, −1 or −2). Mean procedure time was recorded as 22.1 minutes (SD 9.7 minutes).

There was a similar number of procedures evaluated by each of the nine physicians, 15 nurses and four observers. The mean total SPECS scores for the physician, nurse and research assistant were 2.1 (SD 2.0), 2.2 (SD 2.2) and 2.5 (SD 2.2), respectively. The SPECS and GS showed excellent inter-rater reliability among all three raters with an ICC of 0.81 (95% CI, 0.78–0.84) and 0.77 (95% CI, 0.73–0.80), respectively. There was similar inter-rater reliability with SPECS compared with GS for physician versus nurse (ICC 0.78 versus 0.73), nurse versus observer (ICC 0.81 versus 0.75), and physician versus observer (ICC 0.85 versus 0.82). When the individual categories of SPECS were examined, agreement among the three raters was substantial for subject vocalization (weighted κ 0.61; 95% CI, 0.56–0.66) and moderate for position (κ 0.54; 95% CI, 0.48–0.60) and anxiety (κ 0.54; 95% CI, 0.47–0.61).

The mean patient self-reported VAS was 2.2 cm (SD 2.6 cm). Seventy-four percent of patients reported mild pain (VAS ≤ 3 cm), 17% moderate pain, (VAS 3.1 cm to 6.9 cm), and 9% severe pain (VAS ≥ 7 cm). The Spearman correlation between the VAS, the amount of sedation administered, RASS, or procedure time was <0.2, indicating low correlation. The Spearman correlation between VAS and the mean total SPECS score demonstrated moderate positive correlation for the physician (0.52) and observer (0.53) and mild correlation for the nurse (0.42). Among all three raters, the Spearman correlation between subscales of SPECS and VAS was highest for

the vocalization category followed by the positioning and anxiety categories. When the SPECS scores for the anxiety category were compared with the patient-reported anxiety on the post-procedural questionnaire, kappa analysis showed only slight agreement, suggesting limited ability of the raters to predict patient anxiety.

Patient self-reported VAS showed mild to moderate correlation with SPECS ($\rho = 0.53$), GS ($\rho = 0.50$), NPAT ($\rho = 0.47$) and NAPCOMS ($\rho = 0.49$), using the observer's scores for each subject. The mean differences between VAS and the various scores (converted to percentage of maximum possible score) were 4.8% (95% CI, 2.1%–7.6%) for SPECS, 7.8% (95% CI, 5.0%–10.7%) for GS, 3.0% (95% CI, 0.3%–5.7%) for NPAT, and 12.5% (95% CI, 9.6%–15.5%) for NAPCOMS.

Of the 310 (97.8%) subjects who answered the question regarding overall satisfaction on the postprocedure questionnaire, 89.0% were very satisfied with their colonoscopy experience, 6.1% were somewhat satisfied, 3.9% reported fair satisfaction and 1.0% were either somewhat dissatisfied or very dissatisfied. Due to the lack of variability in responses, further statistical analyses were not appropriate.

DISCUSSION

The SPECS is a newly developed patient comfort score assessing verbal and nonverbal cues for pain and anxiety during colonoscopy. The SPECS incorporates measures of both severity and frequency, which may improve standardization of scoring among different health care professionals. The inter-rater validity for SPECS showed excellent reliability among three independent observers. While the SPECS ICC was superior to the commonly used GS, the overlapping confidence intervals indicate this is not a significant improvement. The SPECS correlated moderately well to the patient-reported VAS for comfort and had a similar agreement with the VAS as the other comfort scores assessed.

The SPECS subscales are based on the frequency of patients' behavioural cues of discomfort; therefore, we hypothesized that there would be a direct relationship between SPECS and procedure time. However, this was not observed. This may be due to 74% of patients reporting no or mild pain during colonoscopy despite nearly 90% being mildly sedated. This corresponds to low mean scores on all of the four comfort scales.

The anxiety subclass of SPECS was assessed separately. Observer-assessed patient anxiety did not correlate with patient-reported anxiety. Patients tended to report experiencing higher levels of anxiety during the procedure than the three raters, even when recounted postprocedure. These results suggest that anxiety is more difficult to assess from vocalization and behavioural cues. Health care providers may be able to provide better care for patients by actively inquiring about their anxiety level before the procedure.

Strengths of this study are the prospective design, large sample size, independent recording of three observers and patient blinding. Limitations of this study include possible effects of sedation on the validity of patient-reported outcomes and potential bias in physician's reporting of comfort scores on procedures they have performed. In addition, the distribution of responses for the patient-reported outcomes was narrow, limiting the validity of comparisons to SPECS.

In conclusion, the SPECS is a valid tool for measurement of patient comfort during colonoscopy. The SPECS had excellent inter-rater reliability and correlated moderately well to the VAS. Future studies are needed to improve assessment of anxiety on endoscopic comfort scores.

References

1. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015;64(10):1637–49.
2. Armstrong D, Barkun A, Bridges R, et al.; Canadian Association of Gastroenterology Safety and Quality Indicators in Endoscopy Consensus Group. Canadian Association of Gastroenterology consensus guidelines on safety and quality indicators in endoscopy. Can J Gastroenterol 2012;26(1):17–31.
3. Rex DK, Schoenfeld PS, Cohen J, et al. Quality indicators for colonoscopy. Gastrointest Endosc 2015;81(1):31–53.
4. Ko HH, Zhang H, Telford JJ, et al. Factors influencing patient satisfaction when undergoing endoscopic procedures. Gastrointest Endosc 2009;69(4):883–91, quiz 891.e1.
5. Munson GW, Van Norstrand MD, O'donnell JJ, et al. Intraprocedural evaluation of comfort for sedated outpatient upper endoscopy and colonoscopy: the La Crosse (WI) intra-endoscopy sedation comfort score. Gastroenterol Nurs 2011;34(4):296–301.
6. Rostom A, Ross ED, Dube C, et al. Development and validation of a nurse-assessed patient comfort score for colonoscopy. Gastrointest Endosc 2013;77:255–61.
7. Conjoint Board of the Royal College of Physicians and Royal College of Surgeons Working Group. Guidelines for the Implementation of a National Quality Assurance Programme in GI Endoscopy - Version 2.0. Royal College of Physicians of Ireland,. Last Update, 2011. Available at: https://rcpi-live-cdn.s3.amazonaws.com/wp-content/uploads/2016/01/Guidelines-for-the-Implementation-of-a-National-Quality-Assurance-Programme-in-GI-Endoscopy.pdf. Accessed March 21, 2018.
8. Gélinas C, Puntillo KA, Joffe AM, et al. A validated approach to evaluating psychometric properties of pain assessment tools for use in nonverbal critically ill adults. Semin Respir Crit Care Med 2013;34(2):153–68.
9. Klein DG, Dumpe M, Katz E, et al. Pain assessment in the intensive care unit: development and psychometric testing of the nonverbal pain assessment tool. Heart Lung 2010;39(6):521–8.
10. Ely EW, Truman B, Shintani A, et al. Monitoring sedation status over time in ICU patients: reliability and validity of the Richmond Agitation-Sedation Scale (RASS). JAMA 2003;289(22):2983–91.
11. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8(1):23–34.
12. Cohen P, Cohen J, Aiken LS, West SG. The problem of units and the circumstance for POMP. Multivariate Behavioral Research 1999;34(3):315–346.