Validity of a web-based educational program to disseminate a standardized bowel preparation rating scale

Audrey H. Calderwood, MD[1], **Judith R. Logan, MD, MS**[2], **Michael Zurfluh, MS**[3], **David A. Lieberman, MD**[3], **Brian C. Jacobson, MD, MPH**[1], **Timothy C. Heeren, PhD**[4], and **Paul C. Schroy III, MD, MPH**[1]

[1]Section of Gastroenterology, Boston University Medical Center, Boston, MA

[2]Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, OR

[3]Division of Gastroenterology, Oregon Health & Science University, Portland, OR

[4]Department of Biostatistics, Boston University School of Public Health, Boston, MA

Abstract

Goals—Our goal was to assess the validity of a web-based educational program on the Boston Bowel Preparation Scale (BBPS).

Background—Data on web-based education for improving the practice and quality of colonoscopy are limited.

Study—Endoscopists worldwide participated in the BBPS Educational Program. We assessed program completion rates, satisfaction, short- (0–90 days) and long-term (91–180 days) uptake of the BBPS, and the validity of the program by measuring the reliability of the BBPS among participants.

Results—Two hundred seven endoscopists completed the program. Overall, 93% found the content relevant, 89% felt confident in using the BBPS, and 97% thought the quality was good or excellent. Uptake of the BBPS into clinical practice was robust with 91% and 98% of colonoscopy reports containing the BBPS at short- and long-term follow-up. The interobserver and test-retest reliability of BBPS segment and total scores were both substantial.

Correspondence and reprints: Audrey Calderwood, MD, 85 East Concord Street, Room 7714, Boston, MA 02118, T: 617-638-8211, F: 617-638-6529, audrey.calderwood@bmc.org.

Conclusions—A BBPS web-based educational program facilitates adoption into clinical practice and teaches the BBPS to be used reliably by a diverse group of endoscopists worldwide.

Keywords

education; internet; colonoscopy; bowel preparation; validity

Introduction

The performance of high quality colonoscopy has emerged as an area of critical importance with the surge in demand for colorectal cancer screening and surveillance. Documentation of the cleanliness of bowel preparation has been identified as a quality indicator for colonoscopy and a key element in the colonoscopy report.[1] Poor bowel preparation is associated with lower adenoma detection rates as well as increased time, cost, and need for repeat procedures.[2] Despite the importance of bowel preparation and its impact on colonoscopy outcomes, no single system for evaluating cleanliness has been recommended.[1] The Boston Bowel Preparation Scale (BBPS) is a standardized scoring system for describing bowel cleanliness in the three segments of the colon (right, transverse, and left) after all washing and cleaning maneuvers have been performed and has demonstrated validity and reliability in single-center studies.[3–5] The BBPS has been used for research purposes comparing bowel purgatives[6–10], dietary restrictions[11], and pre-procedure education[12, 13], yet whether it can be adopted widely into clinical practice is unknown.

Web-based education has been used successfully for increasing provider knowledge and is an accepted mechanism for conducting continuing medical education.[14, 15] Currently, the scant data on web-based education specifically for improving the practice and quality of colonoscopy is limited to trainees.[16] The objectives of this study were to develop a web-based educational program to teach the BBPS, evaluate its acceptance among endoscopists worldwide, and monitor uptake of the BBPS in clinical practice. Secondary aims included assessing the ability of the BBPS to be used reliably among those who participate in the online education. We hypothesized that a web-based educational program would be effective at training endoscopists on the use of a standardized bowel preparation scale and that participants would feel confident using the BBPS upon completion of the program and adopt it in their practice.

Materials and Methods

This study was approved by the Institutional Review Board at Boston Medical Center, Boston, MA and Oregon Health Sciences University, Portland, OR.

Development of the educational program

We designed a 20 minute web-based interactive program, The Boston Bowel Preparation Scale Educational Program (BBPSEP), that included the rationale for using a standardized bowel preparation rating scale, a description of the BBPS, and demonstration cases in which the BBPS was used (available online at https://www.cori.org/bbps/login.php). Participants were then asked to assign segment and total BBPS scores during four colonoscopy videos

designed to represent total BBPS scores of 2, 4, 6, and 8. Automated feedback was provided. Participants were required to score 75% or better to complete the program. Continuing medical education (CME) risk management credit of 0.5 hours was offered at no cost. The BBPSEP was beta tested among 9 gastroenterology fellows and 1 attending to assess the content of material, video quality, and time needed to complete the program.

Survey Instrument

A 9-item questionnaire was developed to assess the quality of the BBPSEP, intended change in bowel cleanliness evaluation, and motivation for participation (Appendix 1). The questionnaire was presented at the end of the BBPSEP. Specialty (gastroenterology, family medicine, surgery, other), years in practice (<5, 5–10, 11–20, >20), practice setting (ambulatory endoscopy center, community hospital, academic center, Veterans Affairs, other), annual colonoscopy volume (<100, 100–250, 251–500, 501–1000, >1000), and geographic region (North East, South East, North Central, South Central, North West, South West, outside US including country) were also obtained.

Target Population

Endoscopists worldwide were welcome to participate in the BBPSEP, including those who contacted the investigators directly for information on the BBPS or others who found it independently from the Boston Medical Center Gastroenterology Research page (http://www.bmc.org/gastroenterology/research.htm). In addition, endoscopists affiliated with the Clinical Outcomes Research Initiative (CORI) who perform colonoscopy and use the CORI endoscopy reporting software were invited to participate in the BBPSEP. CORI is a network of practice-based endoscopists selected to represent a complete spectrum of GI practices and to include both urban and rural sites in each region of the country. CORI is composed of mostly community-based private practices (73%) with some academic (13%) and Veterans Affairs (VA) sites (11%).[17]

Outcomes

We evaluated completion rates of the BBPSEP, defined as (# completed/# logged on), satisfaction with the program in terms of perceived relevance and quality, confidence in using the BBPS, and intended use of the BBPS. Among CORI users in whom we had prospective data on the content of their endoscopy reports, short- (0–90 days) and long-term (91–180 days) uptake of the BBPS was calculated as (# of colonoscopies with the BBPS/total colonoscopies performed over that time period).

BBPS Reliability Testing

To evaluate the ability of the BBPSEP to teach the BBPS reliably, we developed a website showing endoscopic still images from 10 different colonoscopies with varying degrees of bowel cleanliness. For each colonoscopy, one image was provided for each segment of the colon (right, transverse, left) for a total of 30 images. Figure 1 shows a selection of the images. The site was beta tested among 3 experts in the BBPS to assess clarity of images. Participants were asked to evaluate the cleanliness of the images using the BBPS on two occasions at least 2 weeks apart, with colonoscopies presented in random order to limit

recall bias. Each viewing was designed to take 3–5 minutes. Our primary endpoint was interobserver reliability of BBPS segment and total scores. Secondary endpoints included test-retest reliability of BBPS segment and total scores.

Analyses

We calculated completion rates of the BBPSEP by calculating the percent of those who logged onto the site that completed the program. Frequencies of responses to survey questions regarding relevance, confidence, and quality were determined as percentages and comparisons were made by CORI status and location (US vs. international) using the chi-square test for independence. Intended use of the BBPS was calculated as percentages and comparisons were made between those who had previously used the BBPS and those who had not using the Fisher's exact test. Among CORI users, we also calculated short- and long-term uptake of the BBPS defined as the percentage of colonoscopies in which a BBPS was included over the first 0–90 days and then 91–180 days, respectively, after completion of the BBPSEP.

We assessed the validity of the BBPSEP to teach the BBPS by measuring how reliably the BBPS was used among those who completed the BBPSEP. The interobserver reliability, the capacity to provide the same measurement among different users, of overall BBPS scores and segment scores as a group and by colon segment was calculated through the intra-class correlation coefficient (ICC(2,1) in the notation of Fleiss).[18] The ICC can be used to measure agreement between more than two raters and is the ratio of between-subject variability to total (between- plus within-subject) variability; the ICC is close to 1.0 when there is little variation in scores by different observers for a subject relative to the variability in scores across different subjects. An ICC below 0.40 represents poor agreement, between 0.4 and 0.75 fair to good agreement, and over 0.75 represents good to excellent agreement.[18]

The test-retest (intra-observer) reliability, the capacity to provide the same measurement on different occasions, was evaluated through the weighted kappa statistic. The kappa statistic describes the percent of agreement beyond chance for two ratings for a categorical rating; for ordinal scaled data, the weighted kappa measures agreement giving a greater penalty for more serious disagreement. A weighted kappa below 0.4 represents poor agreement, between 0.4 and 0.6 moderate agreement, between 0.6 and 0.8 substantial agreement, and over 0.8 almost perfect agreement.[19] All calculations were performed using SAS version 9.3 (SAS Institute, Cary, NC).

Sample size calculation

For the reliability component, we anticipated an ICC for inter-rater reliability above 0.70 based on previous work.[3, 4] Assuming 30 endoscopists participated and each provided 10 overall BBPS scores and 30 BBPS segment scores, a 95% CI for the ICC for BBPS segment scores would have a width of $\pm$ 0.11 and the ICC for the overall BBPS scores would have a width of $\pm$ 0.20.[20]

Results

Sample description

Of the 433 endoscopists who logged onto the BBPSEP website, 207 (48%) completed the program. Ninety-one percent (N=189) of those who completed the program also completed the survey questions. Survey respondents were mostly gastroenterologists working in the community hospital or academic setting with a diversity of experience and annual colonoscopy volume (Table 1). Sixty nine were from outside the United States and 39 were CORI users. Participants affiliated with CORI were in practice longer, less likely to work at an academic hospital, and more likely to practice in the US compared to participants not affiliated with CORI. Among the 114 participants who answered the question on the time it took them to complete the program, the mean (SD) time was 20.3(7.6) minutes.

Among the 226 people who logged in but did not complete the program (defined as scoring 3 out of 4 test cases correctly), 30 (13.3%) did not start the instructional video, 133 (58.5%) watched some portion of the instructional video, and 63 (27.9%) started the test cases (mean number of cases 2.1) but did not complete all 4.

Uptake and Satisfaction

Among the 189 participants who answered the survey questions, there were no significant differences in baseline bowel cleanliness rating method by location of practice in the US or outside the US prior to taking the BBPSEP. Endoscopists unaffiliated with CORI (56%) were less likely to use a qualitative rating scale compared to CORI endoscopists (87%; P=.0001) and were also more likely (29%) to use a dichotomous rating scale compared to CORI endoscopists (0%; P=0.003). Overall, 19% (36/189) of participants said they used the BBPS before the program compared to 84% who said they planned to use the BBPS after completing the BBPSEP (P<.001). Those who had previously used the BBPS (34/36=94.4%) were more likely to plan to use the BBPS in the future compared to those who had never previously used the BBPS (124/153=81.0%) with a trend toward statistical significance (P=.08). International participants (91%) were more likely to intend to use the BBPS than US participants (79%; P=.04).

Among CORI users (N=39), uptake of the BBPS into clinical practice was robust with 74.0% (1928/2606) and 68.4% (1591/2326) of exams containing the BBPS at short-term (0–90 days) and long-term (91–180 days) follow-up, respectively. The median (interquartile range) percentage of colonoscopies with a BBPS score among participants was 91% (34–100) at short-term and 98% (17–100) at long-term follow-up.

Overall, 93% found the program relevant/very relevant, 89% were confident/very confident in using BBPS after completion, and 97% thought the quality was good or excellent. Compared to US participants, international participants were more likely to rate the program as relevant or very relevant (100% vs. 90%1 P=.009). There was no significant difference in other responses by location or CORI status. The primary motivation for participation was the importance of having a validated scale (65%), followed by support of investigators (22%). Only 13% were motivated by CME credit. When asked what could motivate others to participate, recommendations included increasing awareness of the program, department

mandates, involvement in quality improvement projects, and dedicated time at work for participation.

Reliability testing

Reliability testing occurred between December 2012 and March 2013. Of 104 eligible participants at the time, 55 (53%) participated in the reliability study: 40 completed both viewings with a median of 23 (17–28) days between viewings and 15 completed one viewing only. The sample was mostly adult gastroenterologists (80%) working in hospitals (76%) in the community setting (56%). In terms of geographic location, 46% of participants were from outside the US.

The interobserver reliability of segment and total BBPS scores was excellent with ICCs of 0.86 (95% CI 0.74 to 0.93) and 0.90 (95% CI 0.77 to 0.97), respectively.[18] The test-retest reliability of segment and total BBPS scores was also substantial with weighted kappa values of 0.80 (95% CI 0.77 to 0.83) and 0.77 (95% CI 0.74 to 0.80), respectively.[19] The ICCs and weighted kappa values stratified by geographic location, years in practice and annual volume of colonoscopies performed are shown in Table 2. There was some variation in test-retest reliability by segment location, with more reliability of the BBPS for the right and transverse colon compared to the left colon (P <.001) (Table 2C).

Discussion

A short web-based educational program describing the rationale and use of a standardized bowel preparation rating scale was developed and shown to be valid at increasing knowledge and promoting adoption into clinical practice. In addition, the program was effective at teaching those who completed it to use the scale reliably. Our findings provide new evidence supporting the utility of innovative, on-line educational programs such as the BBPSEP for delivering new information to practicing endoscopists and changing behavior.

While simulation-based skills training for endoscopy has been accepted as an adjunct to the "apprenticeship" model,[21] there is limited literature on the use of the Internet for delivery of education to endoscopists. A recent randomized controlled study evaluated the effect of a web-based tutorial on endoscopy-related quality performance measures on gastroenterology trainees.[16] The tutorial, which was comprised of a set of 12 non-interactive questions with answers, modestly improved knowledge in several areas as measured in pre- and post-testing. Whether this knowledge was sustainable or translated into changes in performance was not evaluated. The literature surrounding methods of delivering education to those who have completed formal training is far scarcer and to our knowledge, this is the first study to evaluate the effectiveness of web-based education for the practicing endoscopist.

Previous work on the BBPS was limited to studies that used an educational DVD with colonoscopy footage to educate both trainees and attendings and tested reliability among those at a singlecenteronly.[3, 4] A second validation was performed using the same educational video among a group of attendants at an educational meeting in China.[5] This current work demonstrates that a web-based program can teach the reliable use of the BBPS even among a diverse sample of endoscopists working in various settings across the world.

The observation participants assign segment scores as reliably as total scores may have increasing clinical relevance given concern for variability of bowel preparation quality and missed lesions by colon location. The reason left colon scores had lower test-retest reliability than right and transverse colon scores is unclear. One hypothesis is that the cleanliness of the left colon was more often borderline between two scores (e.g. almost perfectly clean with small spots of debris) and as such could result in users assigning different scores over time (e.g. 2 then 3 or 3 then 2). Even though test-retest reliability of segment scores varied by location, it is reassuring that scores given in the right-side were found to have substantial reliability and that the interobserver reliability of segment scores was high across the entire colon.

Our study has several notable strengths. The diversity of participants from around the world and in diverse practices in the US enhances the external validity of our findings. The quality of the program was rated highly and users were satisfied with it. Importantly, we were able to monitor completion of the educational program, as well as the uptake and adoption of the BBPS into practice over a sustained period of time for a subset of participants. Whereas most studies to date focus on assessment of knowledge alone or confidence after web-based education,[16, 22] relatively few evaluated short term or long term adoption; those who did, relied on self-reported outcomes rather than direct observation.[23] We also demonstrated that the program was able to teach the use of the BBPS reliably.

Our study has limitations. There may be volunteer bias among those who participated. Participants may be from a select group who are already interested in online education and/or quality measures and therefore the feasibility of the program may not apply outside this setting. The program completion rate was 48%. The majority of those who did not complete the program watched the instructional video but did not start any of the test cases, suggesting that perhaps they did not have time to complete or interest in the knowledge assessment piece. We do not have demographic information on those who did not complete the program with which to make comparisons. An abbreviated program of shorter duration might improve completion rates, but whether it would change satisfaction, knowledge, or uptake remains unknown. The CORI reporting software already provides a field for the BBPS and as such may have also facilitated interest in the educational program. We were able to track uptake of the BBPS in real clinical practice within the CORI system but not participants outside of CORI, and therefore can not comment on uptake of the BBPS outside of CORI.

In conclusion, we have demonstrated that a novel web-based educational program is a valid strategy for teaching a diverse audience of practicing endoscopists about the BBPS. Among participants, uptake of the BBPS into practice was robust and its use reliable. The BBPSEP may have increasing utility in clinical practice, especially as the relationship between BBPS scores and outcomes is further defined. The Internet provides an efficient, low cost delivery channel for disseminating the BBPSEP to busy endoscopists in practice.
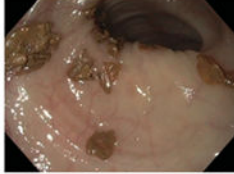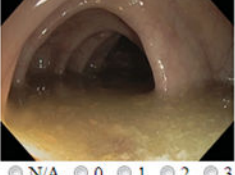
Acknowledgments

References
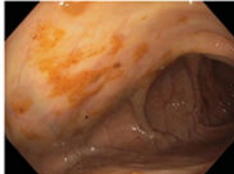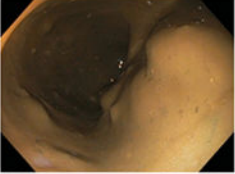
1. Rex DK, Petrini JL, Baron TH, et al. Quality indicators for colonoscopy. Am J Gastroenterol. 2006; 101:873–885. [PubMed: 16635231]

2. Froehlich F, Wietlisbach V, Gonvers JJ, et al. Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study. Gastrointest Endosc. 2005; 61:378–384. [PubMed: 15758907]

3. Calderwood AH, Jacobson BC. Comprehensive validation of the Boston Bowel Preparation Scale. Gastrointest Endosc. 2011; 72:686–692. [PubMed: 20883845]

4. Lai EJ, Calderwood AH, Doros G, et al. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. Gastrointest Endosc. 2009; 69:620–625. [PubMed: 19136102]

5. Gao Y, Lin JS, Zhang HD, et al. Pilot validation of the Boston Bowel Preparation Scale in China. Dig Endosc. 2013; 25:167–173. [PubMed: 23368700]

6. Enestvedt BK, Fennerty MB, Eisen GM. Randomised clinical trial: MiraLAX vs. Golytely - a controlled study of efficacy and patient tolerability in bowel preparation for colonoscopy. Aliment Pharmacol Ther. 2011; 33:33–40. [PubMed: 21083586]

7. Ibanez M, Parra-Blanco A, Zaballa P, et al. Usefulness of an intensive bowel cleansing strategy for repeat colonoscopy after preparation failure. Dis Colon Rectum. 2011; 54:1578–1584. [PubMed: 22067188]

8. McKenna T, Macgill A, Porat G, et al. Colonoscopy preparation: polyethylene glycol with Gatorade is as safe and efficacious as four liters of polyethylene glycol with balanced electrolytes. Dig Dis Sci. 2012; 57:3098–3105. [PubMed: 22711499]

9. Repici A, Cestari R, Annese V, et al. Randomised clinical trial: low-volume bowel preparation for colonoscopy - a comparison between two different PEG-based formulations. Aliment Pharmacol Ther. 2012; 36:717–724. [PubMed: 22924336]

10. Samarasena JB, Muthusamy VR, Jamal MM. Split-dosed MiraLAX/Gatorade is an effective, safe, and tolerable option for bowel preparation in low-risk patients: a randomized controlled study. Am J Gastroenterol. 2012; 107:1036–1042. [PubMed: 22565162]

11. Sipe BW, Fischer M, Baluyut AR, et al. A low-residue diet improved patient satisfaction with split-dose oral sulfate solution without impairing colonic preparation. Gastrointest Endosc. 2012; 77:932–936. [PubMed: 23531424]

12. Calderwood AH, Lai EJ, Fix OK, et al. An endoscopist-blinded, randomized, controlled trial of a simple visual aid to improve bowel preparation for screening colonoscopy. Gastrointest Endosc. 2011; 73:307–314. [PubMed: 21168840]

13. Tae JW, Lee JC, Hong SJ, et al. Impact of patient education with cartoon visual aids on the quality of bowel preparation for colonoscopy. Gastrointest Endosc. 2012; 76:804–811. [PubMed: 22840295]

14. Casebeer LL, Strasser SM, Spettell CM, et al. Designing tailored Web-based instruction to improve practicing physicians' preventive practices. J Med Internet Res. 2003; 5:e20. [PubMed: 14517111]

15. Allison JJ, Kiefe CI, Wall T, et al. Multicomponent Internet continuing medical education to promote chlamydia screening. Am J Prev Med. 2005; 28:285–290. [PubMed: 15766617]

16. Thompson JS, Lebwohl B, Syngal S, et al. Knowledge of quality performance measures associated with endoscopy among gastroenterology trainees and the impact of a web-based intervention. Gastrointest Endosc. 2012; 76:100–106. e1–e4. [PubMed: 22421498]

17. Lieberman DA, Holub J, Eisen G, et al. Utilization of colonoscopy in the United States: results from a national consortium. Gastrointest Endosc. 2005; 62:875–883. [PubMed: 16301030]

18. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979; 86:420–428. [PubMed: 18839484]

19. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]

20. Giraudeau B, Mary JY. Planning a reproducibility study: how many subjects and how many replicates per subject for an expected width of the 95 per cent confidence interval of the intraclass correlation coefficient. Stat Med. 2001; 20:3205–3214. [PubMed: 11746313]

21. Walsh CM, Sherlock ME, Ling SC, et al. Virtual reality simulation training for health professions trainees in gastrointestinal endoscopy. Cochrane Database Syst Rev. 2012; 6:CD008237. [PubMed: 22696375]

22. Cash B, Mitchner NA, Ravyn D. Confidence-based learning CME: overcoming barriers in irritable bowel syndrome with constipation. J Contin Educ Health Prof. 2011; 31:157–164. [PubMed: 21953655]

23. Adams SG, Pitts J, Wynne J, et al. Effect of a primary care continuing education program on clinical practice of chronic obstructive pulmonary disease: translating theory into practice. Mayo Clin Proc. 2012; 87:862–870. [PubMed: 22958990]

Figure 1.
Example Photos from the Boston Bowel Preparation Scale Reliability Study.

Table 1

Characteristics of those who participated in the Boston Bowel Preparation Scale Educational Program (BBPSEP) and answered the survey questions.

Variable	All	CORI	P value
Total, n (%)	189	39	
Specialty			<.0001
Gastroenterology	152 (83)	32 (82)	
Surgery	17 (9)	2 (5)	
Family medicine	5 (3)	5 (13)	
Other	10 (6)	0	
Years in practice			0.0087
<5	66 (60)	6 (33)	
5–10	3 (3)	1 (6)	
11–20	2 (2)	0	
>20	39 (36)	11 (61)	
Practice setting[*]			.06
AEC	27 (15)	13 (33)	
Community Hospital	68 (37)	14 (36)	
Academic	85 (46)	8 (21)	
Veterans Affairs	17 (9)	11 (28)	
Annual colonoscopy volume			NS
<100	31 (17)	5 (14)	
100–250	39 (22)	10 (29)	
251–500	52 (29)	8 (23)	
501–1000	37 (20)	9 (26)	
>1000	23 (13)	3 (9)	
Location			<.0001
United States	120 (63)	38 (97)	
North East	37 (20)	7 (18)	
South East	19 (10)	2 (5)	
North Central	14 (8)	1 (3)	
South Central	3 (2)	2 (5)	
North West	29 (16)	17 (45)	
South West	16 (9)	9 (24)	
Missing	2 (2)	1 (3)	
Outside United States	69 (37)	1 (3)	
Canada	4 (5)		
South America	4(6)		
Europe	49 (71)		

Variable	All	CORI	P value
Middle East	5 (8)		
Asia	2 (3)		
Australia	2 (3)		
Missing	3 (5)		

*Could choose multiple

AEC = ambulatory endoscopy center

BBPSEP = Boston Bowel Preparation Scale Educational Program

CORI = Clinical Outcomes Research Initiative

GI = gastroenterology

HMO = health maintenance organization

PA = physician assistant

Table 2

Reliability of the Boston Bowel Preparation Scale total and segment scores

A. Total Boston Bowel Preparation Scale Scores			
	N	Interobserver reliability ICC (95% CI)	Test-retest reliability Weighted kappa (95% CI)
All	55	0.90 (0.77, 0.97)	0.77 (0.74, 0.80)
Geographic location			
US	29	0.90 (0.72, 0.97)	0.75 (0.71, 0.79)
Outside US	26	0.91 (0.76, 0.97)	0.80 (0.75, 0.84)
Years in practice			
0–10	35	0.90 (0.77, 0.97)	0.79 (0.75, 0.83)
11	20	0.90 (0.63, 0.97)	0.75 (0.70, 0.79)
Annual colonoscopy volume			
0–500	25	0.89 (0.71, 0.97)	0.79 (0.74, 0.83)
501	27	0.91 (0.74, 0.97)	0.76 (0.72, 0.80)

B. Segment Boston Bowel Preparation Scale Scores			
	N	Interobserver reliability ICC (95% CI)	Test-retest reliability Weighted kappa (95% CI)
All	55	0.86 (0.74, 0.93)	0.80 (0.77, 0.83)
Geographic location			
US	29	0.85 (0.64, 0.93)	0.77 (0.73, 0.81)
Outside US	26	0.88 (0.72, 0.94)	0.83 (0.79, 0.86)
Years in practice			
0–10	35	0.87 (0.75, 0.94)	0.82 (0.79, 0.85)
11	20	0.84 (0.46, 0.94)	0.77 (0.73, 0.81)
Annual colonoscopy volume			
0–500	25	0.86 (0.69, 0.93)	0.82 (0.78, 0.86)
501	27	0.85 (0.62, 0.94)	0.78 (0.74, 0.82)

C. By segment location		
	Interobserver reliability ICC (95% CI)	Test-retest reliability Weighted kappa (95% CI)
Right	0.89 (0.66, 0.97)	0.87 (0.84, 0.90)
Transverse	0.84 (0.63, 0.95)	0.84 (0.81, 0.88)
Left	0.80 (0.52, 0.94)	0.57 (0.50, 0.65)

CI = confidence interval

ICC = intra-class correlation coefficient