Published in final edited form as:
*Gastrointest Endosc*. 2010 October ; 72(4): 686–692. doi:10.1016/j.gie.2010.06.068.

# Comprehensive Validation of the Boston Bowel Preparation Scale

**Audrey H. Calderwood, MD** and **Brian C. Jacobson, MD, MPH**
Boston University Medical Center, Boston, MA

## Abstract

**Background—**The Boston Bowel Preparation Scale (BBPS) is a novel bowel cleanliness rating scale that has undergone partial validation previously.

**Objective—**To fully validate the BBPS and assess ease of its dissemination.

**Design—**Observational study.

**Setting—**Various endoscopy units worldwide.

**Subjects—**Endoscopists.

**Interventions—**NA

**Methods—**Video recordings of colonoscopies with varying degrees of cleanliness were viewed twice by gastroenterologists at one medical center. For each video, participants assigned segment and total BBPS scores. Endoscopists worldwide were also surveyed about their experience with the BBPS after viewing an instructional video.

**Main outcome measurements—**Intraclass correlation coefficients (ICC) and weighted kappas assessed inter- and intra-rater reliability, respectively. The BBPS was used among 983 patients undergoing screening colonoscopy.

**Results—**The BBPS demonstrated near-perfect inter-rater reliability (ICC=0.91) and substantial intra-rater reliability (weighted kappa 0.78; 95% CI 0.73–0.84). Among 983 colonoscopies, right and left colon segment scores of 2 or 3 had a multivariate odds ratio of 1.60 (95% confidence interval 1.01–2.55) and 2.58 (95% confidence interval 1.34–4.98) respectively for polyp detection compared to segment scores of 0 or 1. Endoscopists from a variety of settings worldwide found the BBPS easy to implement and applicable to their patient population.

**Limitations—**Single-center reliability testing.

**Conclusions—**The BBPS is a valid and reliable instrument for assessing bowel cleanliness during colonoscopy. Segment scores may represent a standardized way to determine bowel preparation adequacy. The BBPS can be easily disseminated through the use of a brief instructional video.

## INTRODUCTION

The American Society for Gastrointestinal Endoscopy (ASGE) and American College of Gastroenterology (ACG) Taskforce on Quality in Endoscopy suggested colonoscopy reports include an assessment of bowel preparation quality. They proposed use of terms such as “excellent,” “good,” “fair,” and “poor,” despite the lack of standardized definitions.[1] Others have attempted to introduce standardized bowel preparation descriptors, but these have been limited by lack of extensive validation and reliability testing, complexity limiting practical use, or reliance on global assessments that do not address individual colonic segments.[2–4] Given the variability in protection offered by colonoscopy for right- and left-sided advanced neoplasia[5–7], a standardized bowel preparation rating scale that has the ability to capture colon

segment differences may be advantageous. Furthermore, a standardized bowel preparation rating scale should be readily disseminated among endoscopists worldwide.

The Boston Bowel Preparation Scale (BBPS) is a bowel cleanliness rating scale originally designed and validated for use during colonoscopy-oriented research.[8] It relies on the summation of three individual colonic segment scores (from the right, transverse and left colons) to indicate the degree of bowel visualization. Total BBPS scores have been associated with clinical outcomes such as polyp detection rates, recommendations for repeated procedures, and colonoscope insertion and withdrawal times. Reliability testing, however, was based only on three truncated colonoscopy video clips and was limited to the intermediate range of bowel cleanliness.[8] Individual segment scores have not been validated and total BBPS scores have not yet been shown to be reliable across the full spectrum of bowel cleanliness. The current study was designed to assess inter-and intra-rater reliability of total and segment BBPS scores across the full spectrum of possible degrees of bowel cleanliness, to use the BBPS among a diverse patient population, to examine the relationship between individual segment scores and polyp detection rates, and to assess whether the BBPS can be easily disseminated to other medical centers for both clinical and research purposes.

## METHODS

The study was approved by the Institutional Review Board (IRB) of the Boston University Medical Center (BUMC) with waiver of informed consent.

### The Boston Bowel Preparation Scale

The Boston Bowel Preparation Scale (BBPS) has been described in detail previously.[8] Briefly, the BBPS is applied during the *withdrawal* phase of colonoscopy, after all washing, suctioning, and other cleaning maneuvers have been performed by the endoscopist. Good clinical practice dictates that the endoscopist should always clean as much as possible to obtain the highest possible score and ensure the best possible bowel cleanliness. Each of the three segments of the colon (right, including cecum and ascending colon; transverse, including hepatic and splenic flexures; and left, including descending colon, sigmoid and rectum) is given a score from 0–3 defined as follows:

0 = Unprepared colon segment with mucosa not seen due to solid stool that cannot be cleared.

1 = Portion of mucosa of the colon segment seen, but other areas of the colon segment not well seen due to staining, residual stool and/or opaque liquid.

2 = Minor amount of residual staining, small fragments of stool and/or opaque liquid, but mucosa of colon segment seen well.

3 = Entire mucosa of colon segment seen well with no residual staining, small fragments of stool or opaque liquid.

Each of the three segment scores is then summed for a total score of 0–9, where 0 is unprepared and 9 is entirely clean. If an endoscopist aborts a procedure due to an inadequate preparation, then any non-visualized proximal segments are assigned a score of 0. An instructional video demonstrating how to use the BBPS is available online at http://www.bmc.org/gastroenterology/research.htm.

### Assessment of reliability

We recorded 119 colonoscopies at BUMC using DVD Recorders (Multi-Function DVD Recorder VRD-MC3, Sony ®, Tokyo, Japan) connected to our endoscopy unit's video processors (EVIS Excera II CLV-180, Olympus Medical Systems, Tokyo, Japan). Recordings

contained only endoscopic video footage and were free of patient identifiers. From this video library, one author selected ten video clips demonstrating the full range of bowel cleanliness encountered in clinical practice, from excellent to poor, in which no polyps were detected. Each video clip demonstrated the entire withdrawal portion of a colonoscopy, including two cases when the preparation was so poor that only the left sided colon was seen. Using video editing software (Adobe Premiere Elements 3.0, Adobe® Systems Inc., San Jose, CA), we created DVDs, termed bowel preparation DVDs (BP-DVDs), that contained these 10 colonoscopy video clips but in different random orders. The standard bowel preparation in our endoscopy unit at the time the video clips were recorded was a 4 liter polyethylene glycol lavage solution (GoLytely®, NuLytely®, Braintree Laboratories Inc., Braintree, MA, Colyte® or TriLyte® Schwarz Pharma Inc., Milwaukee, WI) taken the night prior to the scheduled colonoscopy. Some patients may have also taken bisacodyl 20 mg prior to beginning their lavage solution. Split dosing was not used during the time of this study.

Attending gastroenterologists and fellows-in-training at BUMC, a tertiary care, academic hospital, viewed all 10 colonoscopies on the DVDs on two occasions at least four weeks apart. Different BP-DVDs that contained videos in different orders were used for the first and second viewings to minimize the physician's ability to remember his or her previous responses (i.e. to limit recall bias). For each video, participants assigned segment and total BBPS scores, categorized the bowel preparation as adequate or inadequate to exclude polyps >5 mm in size, and recommended the timing of the next colonoscopy assuming the colonoscopy was from a 60 year-old healthy man with no family history of colonic neoplasia undergoing routine screening.

### Use of BBPS During Screening Colonoscopy and Relationship to Polyp Detection

After all the endoscopists had viewed the instructional video describing use of the BBPS, they applied the BBPS prospectively during 983 screening colonoscopies at our institution (633 described in our preliminary analysis[8] and 350 additional procedures). The 983 cases were those in which a BBPS score was recorded during an unrelated study of screening colonoscopy (NCT00643682).[9] We excluded patients who had an incomplete procedure for reasons other than bowel preparation (n=7) or had a history of prior colectomy (n=2). After each screening colonoscopy, the endoscopist was asked to record the quality of the bowel preparation using the BBPS. The endoscopists also recorded the location and size of all polyps found during the examinations. We defined diminutive polyps *a priori* as those estimated by the endoscopist to be ≤5 mm.

### Assessment of ability to disseminate the BBPS to other medical centers

Since the publication of our initial BBPS validation study,[8] thirty five endoscopists in numerous countries and practice settings have contacted us requesting access to our BBPS instructional video. This instructional video was made available as a free download from our gastroenterology division's website (as listed above) and the link was sent to the requesting endoscopists. We later invited these endoscopists to complete an IRB-approved, 26-question electronic survey in order to learn about their experience with the instructional video and any subsequent use of the BBPS for either clinical or research purposes (Appendix 1).

### Statistical analysis

To assess inter-rater reliability, we calculated the intraclass correlation coefficient (ICC) (2,1) among total and segment BBPS scores applied after viewing the BP-DVDs, using the methods of Shrout and Fleiss.[10] To assess intra-rater reliability, we calculated weighted kappa measures for total and segment BBPS scores, according to Fleiss and Cohen.[11] We calculated the percentage of physicians deeming a bowel preparation adequate to exclude polyps >5 mm in size for each BBPS score (0–9) assigned during the BP-DVD viewings. The distribution of

continuous variables was tested for normalcy. We then calculated the mean (standard deviation) interval of follow-up to next colonoscopy recommended by participants. Among 983 screening colonoscopies, we compared mean BBPS scores among various demographic categories including gender (male, female), age (<65, ≥65), and self-identified race/ethnicity (White, Black, Hispanic, Asian). In this last comparison, White patients were the reference population. Chi-square tests and analysis of variance were used to compare BBPS scores across groups. Among these 983 colonoscopies, we also determined the polyp detection rate for each colon segment based upon BBPS segment scores. Associations between BBPS scores and polyp detection rates were calculated using Chi-square tests or Fisher's Exact test, and Chi-square tests for trend. Unconditional logistic regression was used to determine the odds ratio for polyp detection controlling for segment scores, age, race and gender. Descriptive statistics were used to analyze survey results. All calculations were performed using SAS version 9.1 (SAS Institute, Cary, NC), and 2-sided *P* values <0.05 were considered significant.

### Sample size estimation

The primary outcomes of this study were the inter- and intra-rater reliability for total BBPS scores. Since the intra-rater reliability is usually higher than or equal to the inter-rater reliability, we performed a power calculation based on expected ICCs to estimate appropriate sample sizes. Based on our preliminary BBPS validation study, we assumed that at least 10 gastroenterologists ("raters") would participate and that the expected ICC would be 0.74 or greater.[8, 12] We then estimated that approximately 6–11 videos ("subjects") would be needed to have a 95% confidence interval with a width of ±0.20 for an ICC between 0.7–0.8.[13] If the ICC approached 0.9, then approximately 8 videos would be needed to have a 95% confidence interval with a width of ±0.10. Given these sample size considerations and the time commitment required by our volunteer raters, we selected ten colonoscopy videos for this study.

A post-hoc analysis showed that we had 75% power to detect meaningful differences in polyp detection for left segment scores of 0 or 1 compared to 2 or 3. We only had 45% power in the right colon to detect such a difference. A total of 5,453 right segment scores (or colonoscopies) would have been needed to achieve 80% power to detect such differences.

## RESULTS

### Reliability testing

The BP-DVDs were viewed by 9 full-time faculty and 3 fellows at BUMC. Previous work has shown that the BBPS can be used reliably by clinicians with various levels of experience with similar inter- and intra-rater reliability[8] and therefore results for all 12 participants were analyzed and reported together. Some of the participants had been introduced to the BBPS during a previous research study[8], but the scale was not being used routinely for rating bowel preparation quality. Individuals viewed a different BP-DVD on two occasions, with a mean (standard deviation (SD)) of 61 ± 35 days between viewings. Table 1 shows the spectrum of bowel cleanliness demonstrated by the ten colonoscopy videos clips in the BP-DVDs.

All physicians deemed the bowel preparation adequate to exclude polyps >5 mm in size when the BBPS score was 8 or greater, compared to 88% when the score was 7, 82% when the score was 6, 33% when the score was 5, and 0% when the score was 4 or less.

When BBPS score was less than 5, all physicians recommended repeat colonoscopy within 1 year. When the BBPS score was between 5 and 6, the mean recommended interval to next colonoscopy was 4.3 (±3.9) years. When BBPS score was 7 or greater, all physicians recommended the next colonoscopy occur in 10 years (Table 1).

The BBPS demonstrated near-perfect inter-rater reliability with an ICC of 0.91 over the full range of possible total BBPS scores. Intra-rater agreement was substantial with a weighted kappa value of 0.78 (95% CI, 0.73–0.84) over the full range of possible total BBPS scores.[14] Inter- and intra-rater reliability for BBPS segment scores by location (right, transverse, left) were also similar (Table 2).

### Use during screening colonoscopy and segment score validation

The BBPS was used prospectively by 12 attending gastroenterologists during 983 screening colonoscopies. The mean (SD) BBPS score was 6.2 ± 1.6, and the median score was 6.0 (range 0.0–9.0; interquartile range 6.0–7.0). Women had slightly higher BBPS scores compared to men (6.4 vs. 6.0, $P < .001$), and Asian patients had slightly higher scores compared to Whites (6.6 vs. 6.2, $P = .05$), but there was no significant difference in mean BBPS scores among other races using Whites as the reference population (Black 6.2, $P = 0.76$; Hispanic 6.1, $P = 0.94$) or by age (6.2 for age <65 vs. 6.3 for age ≥65, $P = 0.42$).

Individual BBPS segment scores in the right colon showed a positive trend with polyp detection rates (Table 3). We found that higher BBPS segment scores (2 and 3 versus 0 and 1) were associated with improved polyp detection in the left colon (OR 2.58, 95%CI 1.34–4.98) and right colon (OR 1.60, 95%CI 1.01–2.55) when controlling for age, gender and race. There was no association found between scores and polyp detection in the transverse colon (OR 0.7, 95% CI 0.48–1.96).

### Ability to Disseminate the BBPS

Twenty four out of 35 (69%) invited participants completed our survey. Of these, sixteen (67%) were adult gastroenterologists, 7 (29%) were gastroenterology fellows, and one (4%) was a pediatric gastroenterology nurse. The number of years in practice of the adult gastroenterologists ranged from <5 years (37%) to 11–20 years (25%) to >20 years (37%). Participants were from the United States (California, Illinois, Minnesota, North Carolina, Ohio, Oregon, Texas, Washington DC), Argentina, Brazil, Israel, Spain, and Singapore. Eighteen (75%) worked in university-based hospitals, 4 (17%) in private practice, and 2 (8%) in military hospitals. All respondents who watched the instructional video felt “very confident” or “confident” in implementing the BBPS. After learning to use the BBPS, 16 (67%) adopted the scale for bowel cleanliness assessment during colonoscopy. Among survey respondents, 21 (88%) had already used the BBPS for rating colonoscopies, with 16 (67%) using it in over 50 cases. Eighteen reported using the BBPS for research purposes, 14 for routine patient care, 3 for teaching and training, and 1 for quality improvement initiatives. Ninety two percent of respondents found the BBPS to be generally applicable to their patient population.

## DISCUSSION

In this study, we have further established the validity and reliability of the BBPS as an instrument for assessing bowel cleanliness during colonoscopy for both research and clinical purposes. We have also demonstrated that the BBPS can be easily disseminated to a wide range of practice locations worldwide using a simple, online video. Total BBPS scores demonstrated strong inter- and intra-rater reliability over the full range of possible scores and bowel cleanliness. Total BBPS scores were associated with perception of bowel preparation adequacy for excluding polyps >5 mm in size and recommended timing of repeat colonoscopy, both of which provide additional evidence of the construct validity of the BBPS. Individual segment BBPS scores also demonstrated strong inter- and intra-rater reliability over the full range of possible segment scores. Prospective use of the BBPS during more than 900 screening colonoscopies demonstrated significant associations between individual segment scores and polyp detection rates. Gastroenterologists and gastroenterology fellows throughout several

countries and with varying degrees of experience all felt confident in using the BBPS after viewing a short instructional video. Almost all respondents found the BBPS to be generally applicable to their patient population.

Dichotomized BBPS segment scores of 2 and 3, compared to 0 and 1, were associated with an increased likelihood of finding polyps in the right and left colon. Surprisingly, no association was found between transverse colon segment scores and polyp detection. Fewer absolute polyps were detected in the transverse colon as compared to the right and left colons, perhaps limiting the power to detect meaningful differences.

The lack of association between transverse colon segment scores and polyp detection raises the possibility of a simplified BBPS that only includes assessments of the right and left colon. While future research may indeed justify exclusion of the transverse colon segment score, we feel it is likely premature to draw this conclusion as there may be important information captured by inclusion of all three segment scores for research or clinical purposes.

Prospective use of the BBPS during over 900 colonoscopies showed small variations in quality of bowel cleanliness by gender and race. Women had slightly cleaner preparations compared to men, a finding consistent with other studies.[15–17] In our study, we also found that Asians had cleaner preparations compared to Whites, a trend seen in one previous study.[15] This information may be relevant for those planning to use the BBPS for research purposes in special populations, although in general, the differences observed in mean scores were small and likely of limited clinical significance.

Expanding further on our previous validation of total BBPS scores, we have now shown that the BBPS is sensitive to differences in cleanliness within the colonic segments, and therefore may become a useful tool in helping to understand segment-specific risks for missed pathology based on the degree of bowel cleanliness. This feature may prove clinically relevant given reported discrepancies in the effectiveness of colonoscopy for preventing colorectal cancer of the left versus the right colon.[5–7] Such discrepancies may reflect location-specific differences in bowel preparation adequacy and BBPS segment scores may highlight where endoscopists should focus more time and effort in cleaning. Through its ability to reflect the clinical effectiveness of colonoscopy and to preserve variation in cleanliness by segment, the BBPS may be a useful tool in quality assurance for colonoscopy. The BBPS can be presented as a total score (e.g. 6), individual segment scores (e.g. 1-3-2), or both (e.g. 1-3-2=6) to fit the user's needs, recognizing that only individual segment scores preserve segmental differences in bowel preparation quality.

Our study has several strengths including assessment of both inter- and intra-rater reliability of total and segment scores, correlation of BBPS score with the perception of bowel preparation adequacy in excluding polyps >5 mm in size, and prospective validation among a large number of ethnically diverse patients. Nevertheless, we acknowledge certain limitations. The BBPS requires assessment of bowel preparation quality in the right, transverse, and left colon segments and landmarks distinguishing these segments may be poorly defined or difficult to recognize, raising the possibility of difficulty identifying specific segments. While this has the potential to result in assignment of segment scores that do not reflect the intended location, this limitation is not unique to the BBPS, but rather inherent to the practice of colonoscopy and affects other factors, such as the reporting of polyp location. However, we also hypothesize that the excellent inter-rater reliabilities observed for various segments indicates at least reasonable agreement among endoscopists about where the colonoscopes were during score assignments.

Reliability testing was performed using only the withdrawal portion of ten colonoscopies and did not include visualization during colonoscope insertion. However, as the BBPS is applied

only during the withdrawal portion of a colonoscopy, our study methods represent real-world use of the scale.

Our study was limited to a single, urban, academic medical center that serves a diverse, underserved, predominantly Black population (39% Black, 23% White and 20% Hispanic), and therefore our findings may not be generalizable. However, the BBPS performed very similarly among a heterogeneous group of patients, suggesting that it should be useful in various clinical settings. Furthermore, a broad range of endoscopists from multiple national and international settings found the BBPS applicable to their patient populations. The BBPS was also not evaluated for use during other bowel imaging procedures, such as CT colonography or capsule endoscopy, or in colonoscopies after colorectal surgery, and as such, we are unable to comment on its utility in these settings. Until the BBPS is validated in patients who have undergone colorectal surgery or whose procedures are aborted for reasons other than bowel preparation quality, we would suggest that segment scores, rather than total BBPS scores, should be used. While the BBPS has not been validated specifically for use during colonoscopy with the water infusion technique[18], it should be applicable in this setting as clear fluid that does not obscure visualization of the underlying colonic mucosa does not impact the BBPS scoring system. Similarly, if an area of mucosa is initially obscured by overlying residue that is shifted away by a change in patient position such that the underlying mucosa is now well visualized, then the BBPS score should reflect this improvement in visualization.

In summary, the BBPS is a valid and reliable instrument for assessing bowel preparation adequacy during colonoscopy regardless of degree of cleanliness. It may be useful during colonoscopy-related research when needing to control for bowel preparation quality. Furthermore, segment scores may represent a new way to determine bowel preparation adequacy in clinical practice. The BBPS can also be disseminated easily through use of a brief, instructional, web-based video, making it a strong candidate for a standardized tool for rating bowel preparations.

## Acknowledgments

We would like to acknowledge the efforts of our colleagues who reviewed the BP-DVDs as well as those national and international clinicians who responded to our survey. This study was made possible by an ASGE Endoscopic Research Award and a National Institute of Diabetes and Digestive and Kidney Diseases Career Development Award (DK-070706) to BCJ.

## References

1. Rex DK, Petrini JL, Baron TH, et al. Quality indicators for colonoscopy. Am J Gastroenterol 2006;101:873–85. [PubMed: 16635231]
2. Aronchick CA, Lipshutz WH. Validation of an instrument to assess colon cleansing [abstract]. Am J Gastroenterol 1999;94:2667.
3. Aronchick CA, Lipshutz WH, Wright SH, Dufrayne F, Bergman G. A novel tableted purgative for colonoscopic preparation: efficacy and safety comparisons with Colyte and Fleet Phospho-Soda. Gastrointest Endosc 2000;52:346–52. [PubMed: 10968848]
4. Rostom A, Jolicoeur E. Validation of a new scale for the assessment of bowel preparation quality. Gastrointest Endosc 2004;59:482–6. [PubMed: 15044882]
5. Brenner H, Hoffmeister M, Arndt V, Stegmaier C, Altenhofen L, Haug U. Protection from right- and left-sided colorectal neoplasms after colonoscopy: population-based study. J Natl Cancer Inst 2010;102:89–95. [PubMed: 20042716]
6. Lakoff J, Paszat LF, Saskin R, Rabeneck L. Risk of developing proximal versus distal colorectal cancer after a negative colonoscopy: a population-based study. Clin Gastroenterol Hepatol 2008;6:1117–21. quiz 064. [PubMed: 18691942]

7. Baxter NN, Goldwasser MA, Paszat LF, Saskin R, Urbach DR, Rabeneck L. Association of colonoscopy and death from colorectal cancer. Ann Intern Med 2009;150:1–8. [PubMed: 19075198]
8. Lai EJ, Calderwood AH, Doros G, Fix OK, Jacobson BC. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. Gastrointest Endosc 2009;69:620–5. [PubMed: 19136102]
9. Lai EJ, Oren OK, Jacobson BC. A Single-Blinded, Randomized, Controlled Trial Using a Simple Visual Aid to Improve Bowel Preparation Prior to Screening Colonoscopy. Gastrointest Endosc 2007;65:AB328.
10. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420–8. [PubMed: 18839484]
11. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Edu Psychol Meas 1973;33:613–9.
12. Karanicolas PJ, Bhandari M, Kreder H, et al. Evaluating agreement: conducting a reliability study. J Bone Joint Surg Am 2009;91 (Suppl 3):99–106. [PubMed: 19411507]
13. 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;20:3205–14. [PubMed: 11746313]
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74. [PubMed: 843571]
15. Lebwohl B, Wang TC, Neugut AI. Socioeconomic and Other Predictors of Colonoscopy Preparation Quality. Dig Dis Sci. [epub 2010 Jan 16 ahead of print].
16. Borg BB, Gupta NK, Zuckerman GR, Banerjee B, Gyawali CP. Impact of obesity on bowel preparation for colonoscopy. Clin Gastroenterol Hepatol 2009;7:670–5. [PubMed: 19245852]
17. Ness RM, Manam R, Hoen H, Chalasani N. Predictors of inadequate bowel preparation for colonoscopy. Am J Gastroenterol 2001;96:1797–802. [PubMed: 11419832]
18. Leung JW, Mann SK, Siao-Salera R, et al. A randomized, controlled comparison of warm water infusion in lieu of air insufflation versus air insufflation for aiding colonoscopy insertion in sedated patients undergoing colorectal cancer screening and surveillance. Gastrointest Endosc 2009;70:505–10. [PubMed: 19555938]

**Table 1**

Total Boston Bowel Preparation Scale scores of ten colonoscopy videos viewed by 12 gastroenterologists.

| | Video A | Video B | Video C | Video D | Video E | Video F | Video G | Video H | Video I | Video J |
|---|---|---|---|---|---|---|---|---|---|---|
| Total BBPS score, mean (SD)* | 0 (0) | 1.7 (0.9) | 1.9 (0.7) | 3.1 (0.8) | 5.0 (1.2) | 5.2 (0.5) | 6.3 (0.9) | 7.2 (0.8) | 7.3 (1.2) | 8.0 (1.1) |
| % Saying preparation was adequate to exclude polyps >5mm | 0 | 0 | 0 | 0 | 33 | 25 | 75 | 100 | 100 | 100 |
| Recommended screening interval in years, mean (SD) | 0 (0) | 0 (0) | 0 (0) | 0.3 (0.9) | 3.3 (4.9) | 2.3 (3.9) | 7.3 (3.9) | 10 (0) | 10 (0) | 10 (0) |

* Mean of 12 raters assigning a BBPS score to the same video clips; BBPS = Boston Bowel Preparation Scale; SD= standard deviation

## Table 2

Intraclass correlation coefficients (ICC) and weighted kappas for total BBPS scores and segment scores by location demonstrated consistent inter-rater and intra-rater reliability, respectively.

| | Total score | Right colon score | Transverse colon score | Left colon score |
|---|---|---|---|---|
| **ICC** | 0.91 | 0.88 | 0.83 | 0.79 |
| **Weighted kappa (95% CI)** | 0.78 (0.73–0.84) | 0.78 (0.71–0.85) | 0.73 (0.64–0.81) | 0.75 (0.67–0.83) |

ICC = Intraclass correlation coefficient; CI = confidence interval; BBPS = Boston Bowel Preparation Scale

## Table 3

Polyp detection rates for BBPS segment scores by segment location.

| | BBPS segment score | | | |
|---|---|---|---|---|
| | **0** | **1** | **2** | **3** |
| Right colon | 1/35 (3%) | 24/194 (12%) | 91/617 (15%) | 27/137 (20%) |
| Transverse colon | 0/17 (0%) | 10/78 (13%) | 61/643 (9%) | 22/245 (9%) |
| Left colon | 0/7 (0%) | 11/79 (14%) | 158/582 (27%) | 70/315 (22%) |

BBPS = Boston Bowel Preparation Scale