

Performance of bowel preparation quality scales in patients with Crohn's disease

Virginia Solitano^{1,2,3} | Corey A. Siegel⁴ | Joshua R. Korzenik^{5,6} | Jennifer K. Maratt^{7,8}  | Douglas K. Rex⁷ | Bryan Maguire² | Brian Bressler⁹ | Johannes Grossmann¹⁰ | Rocio Sedano^{1,2,11} | John W. D. McDonald^{1,2} | Julie Remillard² | Lisa M. Shackelton² | Guangyong Zou^{2,11} | Brian G. Feagan^{1,2,11} | Christopher Ma^{2,12,13}  | Vipul Jairath^{1,2,11} 

¹Department of Medicine, Division of Gastroenterology, Western University, London, Ontario, Canada

²Alimentiv Inc, London, Ontario, Canada

³Division of Gastroenterology and Gastrointestinal Endoscopy, IRCCS Ospedale San Raffaele, Università Vita-Salute San Raffaele, Milan, Italy

⁴Section of Gastroenterology and Hepatology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA

⁵Division of Gastroenterology, Hepatology and Endoscopy, Brigham and Women's Hospital, Boston, Massachusetts, USA

⁶Harvard Medical School, Boston, Massachusetts, USA

⁷Division of Gastroenterology and Hepatology, Indiana University School of Medicine, Indianapolis, Indiana, USA

⁸Richard L. Roudebush Veterans Affairs Medical Center, Indianapolis, Indiana, USA

⁹Department of Medicine, Division of Gastroenterology, University of British Columbia, Vancouver, Canada

¹⁰RKM740-Interdisciplinary Specialty-Clinic, Gastroenterology, Düsseldorf, Germany

¹¹Department of Epidemiology and Biostatistics, Western University, London, Ontario, Canada

¹²Division of Gastroenterology and Hepatology, Department of Medicine, University of Calgary, Calgary, Alberta, Canada

¹³Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada

Correspondence

Vipul Jairath, Department of Medicine,
Division of Gastroenterology, Department
of Epidemiology and Biostatistics, Western
University, Suite 200, 100 Dundas Street,
London ON N6A 5B6, Canada.
Email: vjairath@uwo.ca

Funding information

Leona M. and Harry B. Helmsley Charitable
Trust, Grant/Award Number: #2005-03895

Summary

Background: The performance of bowel preparation (BP) in patients with Crohn's disease (CD) is unknown.

Aims: To evaluate the operating properties of instruments used to assess BP quality in patients with CD.

Methods: We used the Boston Bowel Preparation Scale, modified Boston Bowel Preparation Scale, Harefield Cleansing Scale, Food and Drug Administration Bowel Cleansing Assessment Scale (BCAS), and a 100-mm visual analogue scale of bowel cleanliness to assess BP quality in 50 videos from 40 patients with CD. We assessed endoscopic activity with the Simple Endoscopic Score for CD (SES-CD). Assessments were on endoscope insertion and withdrawal. Reliability was quantified using the intraclass correlation coefficient (ICC). We assessed validity by within-patient correlation between instruments and the visual analogue scale using mixed-effect models. The correlation between BP quality and SES-SD scores was assessed using Spearman's rho.

The Handling Editor for this article was Professor Cynthia Seow, and it was accepted for publication after full peer-review.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Alimentary Pharmacology & Therapeutics* published by John Wiley & Sons Ltd.

Results: Inter- and intra-rater reliability for all BP quality instruments was substantial ($ICC \geq 0.61$) except for the Food and Drug Administration BCAS on insertion (inter-rater reliability $ICC \geq 0.41$). The visual analogue scale had substantial inter- and almost perfect ($ICC \geq 0.81$) intra-rater reliability. Correlation coefficients for the validity of the instruments exceeded 0.58. BP quality and endoscopic disease activity scores in the colon were negatively correlated.

Conclusion: Most existing instruments reliably assess BP quality in patients with CD. These results support the use of these instruments in clinical practice, provide a framework for scoring BP quality in CD clinical trials, and support evaluation of novel BP agents in patients with CD.

1 | INTRODUCTION

High-quality bowel preparation (BP) is essential for colorectal cancer screening. Risks of suboptimal BP include lower adenoma detection and caecal intubation rates, and increased costs and inconvenience for repeated exams.^{1–3} Adequacy of BP is a reporting standard and quality metric for screening and surveillance colonoscopy.^{4,5} Patients with inflammatory bowel disease (IBD) represent a population of special interest due to the increased lifetime colonoscopy burden for disease activity and response to therapy assessments and dysplasia surveillance.^{6,7} Patients with IBD report decreased tolerability and higher burden with BP, and the quality of bowel cleansing might be impaired by active inflammation or other disease-related factors, such as altered anatomy, including stricturing disease.^{8–10}

Although endoscopic disease activity is formally assessed by central readers using validated instruments in IBD clinical trials,¹¹ BP quality is not rigorously assessed in the same manner. Furthermore, few clinical trials have evaluated BP efficacy in patients with IBD.^{8,12,13} Therefore, the performance of BP-quality instruments has not been well evaluated in patients with IBD. A minimum BP quality standard for adequate examination of disease activity and/or dysplasia surveillance has also not been established.

Existing instruments for BP quality assessment in the general population vary according to the number of colonic segments scored, timing for scoring relative to washing, and assessment of quantity of washing or suctioning required for optimal visualisation.¹⁴ The inter-rater reliability of five BP quality instruments (the Boston BP Scale [BBPS], the Chicago BP scale, the Harefield Cleansing Scale [HCS], the Ottawa BP Quality Scale, and the Aronchick score) and a visual analogue scale (VAS) of BP quality ranged from “moderate” to “substantial” when BP quality was assessed by central readers using videos from two clinical trials including healthy volunteers (NCT02254486 and NCT02273141). Reliability was similar when BP quality was based on assessment of the overall procedure (insertion/withdrawal) and during the withdrawal phase alone.¹⁵ A recent United States Food and Drug Administration (US FDA)¹⁶ draft guidance on efficacy and safety considerations for clinical development of bowel cleansing products for colonoscopy recommended use of a 4-point qualitative instrument

(e.g., the Bowel Cleansing Assessment Scale [BCAS]), assessment of BP quality within individual bowel segments by at least one central reader, and evaluation of BP efficacy on colonoscope insertion (before washing and suctioning) to avoid potential bias related to intraprocedural cleansing. Additionally, a modified Research and Development/University of California, Los Angeles appropriateness methodology (RAM) study determined that central and local readers should be trained on standardised methods for BP quality assessment in Crohn’s disease (CD) clinical trials, and that local endoscopists should additionally be trained on methods for high-quality video recording. In this study, expert panellists also concluded that BP quality should be assessed on both endoscope insertion and withdrawal in all colonic segments, consistent with segmental scoring of CD endoscopic disease activity, and that the BBPS and HCS were appropriate instruments for assessment of BP in clinical trials involving patients with CD.¹⁷

As the performance of BP quality instruments in patients with CD is unknown, we aimed to evaluate the operating properties of appropriate BP instruments and novel items and approaches identified in our prior study,¹⁷ as well as the FDA BCAS when used by central readers for assessment of BP quality in patients with CD.

2 | METHODS

The primary objective was to explore the operating properties of the BBPS, a modified version of the BBPS (mBBPS), HCS, and FDA BCAS in patients with CD, with an overall aim of determining the optimal instrument for BP quality assessment in clinical trials including patients with CD.

2.1 | Study population

Fifty complete (full insertion and withdrawal phases) videos from 40 patients with CD who underwent endoscopy as a component of routine patient care for the assessment of disease activity and which were archived in the University of California San Diego Research Biobank (San Diego, CA) were selected for inclusion. The videos

were identified by a trained endoscopist (RS; who did not participate as a central reader) with the aim of ensuring representation of a range of BP quality and endoscopic disease activity, and to include videos from patients with normal anatomy and with prior intestinal resection.

2.2 | Central reading

Three central readers independently evaluated BP quality and endoscopic disease activity in the ileum, right colon, transverse colon, left colon, and rectum in all videos, on two occasions, separated by at least 2 weeks to facilitate memory extinction. The quality of BP was assessed (twice) on both endoscope insertion and withdrawal whereas endoscopic disease activity was assessed (twice) only on endoscope withdrawal. Central readers were a priori trained on the use of Alimentiv's Central Image Management Solutions for the assessment of videos. Standardised training materials on the scoring of the BP indices (BBPS, mBBPS, HCS, FDA BCAS; see [Material S1](#) for a description of the instruments), the SES-CD, a 100mm VAS of BP quality, and a 100mm VAS of endoscopic severity were additionally provided to central readers.

2.3 | Assessment of BP quality

Bowel preparation quality was assessed with the BBPS, mBBPS, HCS, FDA BCAS, and a 100-mm VAS of BP quality/cleanliness.

The BBPS assesses BP on a scale from 0 to 3 in the right colon (including the cecum and ascending colon), transverse colon (including the hepatic and splenic flexures), and left colon (including the descending and sigmoid colon and rectum).^{18,19} Total BBPS scores range from 0 to 9 and are calculated as the sum of the scores from each colonic region. Higher scores indicate superior cleanliness ([Table S1](#)). Modifications to the BBPS (e.g., the mBBPS) for the assessment of BP in patients with CD based on the prior RAM study¹⁷ included quantifying the amount of staining, residual stool, and/or opaque liquid as covering ≤10% or >10% of mucosa for BBPS scores of 1 or 2. The HCS is a 5-point instrument that assesses BP in 5 colonic segments (ascending colon and cecum, transverse colon, descending colon, sigmoid colon, and rectum) after washing or suctioning. Scores range from 0 to 4, with higher scores indicating superior cleanliness.²⁰ Assessments for bowel segments are expressed as four grades (A–D), where A and B denote successful cleansing ([Table S2](#)). The FDA BCAS is a qualitative instrument that rates BP as excellent, good, fair, and poor ([Table S3](#)). The VAS of BP quality/cleanliness ranged from 0 to 100 where a score of 0 represented "cleanest" and a score of 100 represented the "worst ever seen."

2.4 | Assessment of endoscopic disease activity

Endoscopic disease activity was assessed using the Simple Endoscopic Score for Crohn's Disease (SES-CD) and a 100-mm

VAS of endoscopic disease severity. The SES-CD is comprised of four items (size of mucosal ulcers, degree of ulcerated surface, endoscopic extension, and presence of stenosis) scored on a 4-point scale (0–3) in the ileum, right colon, transverse colon, left colon, and rectum ([Table S4](#)).²¹ Total SES-CD scores range from 0 to 56, calculated as the sum of the segmental scores. Higher scores indicate more severe endoscopic disease activity. The VAS of endoscopic disease severity ranged from 0 to 100 where a score of 0 represented "completely normal" and a score of 100 represented the "worst ever seen."

2.5 | Statistical methods

The reliability of the BP instruments was evaluated with the intra-class correlation coefficient (ICC). Point estimates for ICCs were determined separately for endoscope insertion and withdrawal using a two-way random effects analysis of variance model with interaction, as described by Gilder et al.²² Associated two-sided 95% confidence intervals (CIs) were obtained using non-parametric cluster bootstrapping, with 1000 samples resampled at the level of the video to maintain data structure. These ICCs are equivalent to the weighted Kappa in the case of ordinal data with quadratic weights. The degree of reliability was interpreted according to benchmarks proposed by Landis and Koch where ICCs <0.00 indicate "poor" reliability, and those ranging from 0.0 to 0.20, 0.21–0.40, 0.41–0.60, 0.61–0.80, and 0.81–1.00 indicate "slight", "fair", "moderate", "substantial", and "almost perfect" reliability.²³ Inter-rater ICCs for reliability of the total scores for each BP instrument were compared when assessed on endoscope insertion and withdrawal using the CI approach for comparing two dependent ICCs described by Kowalik et al.²⁴

The validity of the BP instruments was evaluated by within-patient correlations between each instrument and the VAS of bowel cleanliness across segments during endoscope insertion and withdrawal based on quality scores from all readers. Point estimates for within-patient correlations were determined using mixed-effect models.^{25,26} Associated two-sided 95% CIs were constructed using cluster bootstrap methods with 1000 replications.²⁷

The correlation between endoscopic disease activity and BP quality was explored within bowel segments using SES-CD and BP quality scores determined by the first central reader. Associations were quantified using Spearman's rank correlation coefficients with associated two-sided 95% CIs. Correlations were interpreted using Cohen's effect sizes (coefficients equal to 0.1 were considered small correlations, those equal to 0.3 were considered moderate correlations, and those equal to 0.5 were considered large correlations).²⁸

Sample size calculation was based on the one-way random effects model using methods proposed by Zou.²⁹ Assuming a true ICC of 0.80 scoring 50 videos by three central readers would yield a greater than 86% chance of obtaining a one-sided 95% lower bound that was greater than 0.65, corresponding to "substantial" agreement based upon Landis and Koch benchmarks.²³ This estimate was conservative since the data were analysed using a

two-way random effects model that is more efficient than a one-way model.

2.6 | Ethical considerations

The study was conducted in compliance with the protocol, the International Council for Harmonisation (ICH) Good Clinical Practice (GCP), and applicable regulatory requirements. Institutional Review Board approval for the study was granted on 18 February 2022, along with a waiver of consent as consent for the use of endoscopy videos for research purposes had previously been granted by all patients whose videos were selected for this study.

3 | RESULTS

3.1 | Demographic and baseline characteristics

The 50 endoscopy videos assessed in this study included recordings of 34 colonoscopies including terminal ileal evaluation, 14 co-colonoscopies, and 2 flexible sigmoidoscopies performed in 40 CD patients. Demographic data were available for 39 patients (Table 1), of whom 25 (64.1%) were female. Most patients whose videos were assessed were Caucasian (31 [79.5%]). Median age at CD diagnosis was 21 (range, 12–77). Disease location at diagnosis was ileo-colonic in 15 [38.5%], colonic in 6 [15.4%], and ileal in 5 [12.8%] patients (data were missing for 13 [33.3%] patients). Disease behaviour at diagnosis was non-stricturing/non-penetrating in 18 [46.1%], penetrating in 8 [20.5%], and stricturing in 3 [7.7%] patients (data were missing for 10 [25.6%] patients). When data were analysed at the video-level, 14 (28%) videos were from patients with a history of prior surgery and 36 (72%) were from patients with no history of prior surgery. Median baseline SES-CD score based on video-level assessment was 5.5 (range, 0–24), and 27.5% (11/40) of the videos had SES-CD scores ≤ 3 .

3.2 | Reliability of instruments for assessment of BP quality

Substantial ($ICC \geq 0.61$) inter-rater reliability was observed for all BP instruments during endoscope insertion and withdrawal, except for the FDA BCAS, which had moderate ($ICC \geq 0.41$) inter-rater reliability on insertion (Table 2). Point estimates for ICCs for inter-rater reliability of the instruments when assessed on insertion ranged from 0.51 (95% CI 0.32, 0.65; FDA BCAS) to 0.69 (95% CI 0.49, 0.80; mBBP) and from 0.62 (95% CI 0.42, 0.74; FDA BCAS) to 0.72 (95% CI 0.51, 0.81; mBBP) when assessed on withdrawal. Although point estimates for inter-rater reliability were numerically higher for all instruments in the withdrawal phase, these marginal differences between the ICC

estimates during insertion and withdrawal were not statistically significantly different. Intra-rater reliability was similarly substantial for all instruments and almost perfect ($ICC \geq 0.81$) for the VAS during insertion and withdrawal. Point estimates for ICCs for intra-rater reliability of the instruments when assessed on insertion ranged from 0.69 (95% CI 0.53, 0.81; BBPS) to 0.80 (95% CI 0.65, 0.88; HCS) and from 0.74 (95% CI 0.57, 0.82; mBBPS) to 0.79 (95% CI 0.67, 0.85; FDA BCAS) on withdrawal (although ICCs for intra-rater reliability of the BBPS when assessed on withdrawal were nearly identical to those of the mBBPS, as were ICC estimates for the HCS compared to the FDA BCAS). Marginal differences between the ICC estimates for intra-rater reliability of the instruments when assessed during insertion and withdrawal were not statistically significantly different.

3.3 | Validity of instruments for assessment of BP quality

Large correlations were observed between all instruments evaluated for the assessment of BP and the VAS of BP quality/cleanliness (Table 3). Correlation coefficients exceeded 0.58 on both insertion and withdrawal. Although the correlation coefficients were somewhat numerically higher when determined for the withdrawal phase of the procedure for all instruments, the 95% CIs for the correlation coefficients overlapped and no significant differences for correlation with the VAS of BP quality/cleanliness were observed between the instruments or according to the phase of the procedure.

At least a moderate correlation was also observed between the instruments and the VAS of BP quality/cleanliness in all colonic segments during the withdrawal phase (Table 4). Correlation between the instruments and the VAS was consistently lowest in the left colon (coefficients ranging from 0.51 to 0.57), and highest in the right colon (coefficients ranging from 0.62 to 0.73). Whereas correlation coefficients exceeded 0.51 in all colonic segments, they ranged from 0.39 to 0.46 in the ileum.

3.4 | Association between BP quality and endoscopic disease activity scores

Small-to-moderate correlations between BP quality and endoscopic disease activity as measured by the SES-CD were observed in all segments assessed, with negative correlations between BP quality and endoscopic disease activity observed in the colon for all instruments (coefficients ranging from -0.05 to -0.33) compared to positive correlations in the ileum (coefficients ranging from 0.08 to 0.25) (Table 5). Correlation coefficients were consistently higher in the transverse and left colon for all instruments, with small-to-moderate negative correlations (coefficients ranging from -0.26 to -0.33) observed for all instruments in these locations except for the HCS (coefficients ranging from -0.16 to -0.18).

TABLE 1 Patient- and video-level demographic and disease-related variables.

Characteristic	Patient-level (N=39 ^a)	Video-level (N=49 ^a)
Sex, n (%)		
Female	25 (64.1%)	30 (61.2%)
Male	14 (35.9%)	19 (38.8%)
Ethnicity, n (%)		
African American	2 (5.1%)	2 (4.1%)
Asian	1 (2.6%)	1 (2.0%)
Caucasian	31 (79.5%)	40 (81.6%)
Hispanic or Latino	2 (5.1%)	2 (4.1%)
Jewish	1 (2.6%)	2 (4.1%)
Middle Eastern	2 (5.1%)	2 (4.1%)
Tobacco use, n (%)		
Current	1 (2.6%)	1 (2.1%)
Former	6 (15.4%)	8 (16.3%)
Never	32 (82.0%)	40 (81.6%)
Median age at diagnosis, years (min, max)	21 (12, 77)	23 (12, 77)
Location of Crohn's disease at diagnosis, n (%)		
L1: Ileal	5 (12.8%)	7 (14.3%)
L2: Colonic	6 (15.4%)	8 (16.3%)
L3: Ileocolonic	15 (38.5%)	16 (32.7%)
Not available	13 (33.3%)	18 (36.7%)
Behaviour of Crohn's disease at diagnosis, n (%)		
B1: Non-stricturing, non-penetrating disease	18 (46.1%)	23 (46.9%)
B2: Stricturing disease	3 (7.7%)	5 (10.2%)
B3: Penetrating disease	8 (20.5%)	8 (16.3%)
Not available	10 (25.6%)	13 (26.5%)
Procedure, n (%)		
Endoscopy		14 (28%)
Flexible sigmoidoscopy		2 (4%)
Ileocolonoscopy		34 (68%)
Any prior surgery, n (%)		
No		36 (72%)
Yes		14 (28%)
Number of prior surgeries, n (%)		
0		36 (72%)
1		9 (18%)
2		3 (6%)
5		2 (4%)
Median time since diagnosis, months (min, max)		75.2 (7.86, 552.3)
Median SES-CD score, (min, max)		5.5 (0, 24)
SES-CD score ≤3, n (%)		11 (27.5)

Abbreviation: SES-CD, simple endoscopic score for Crohn's disease.

^aData are missing for one patient whose video was included for central assessment.

4 | DISCUSSION

The primary objective of bowel cleansing is the complete removal of faecal matter and fluids from the colon to ensure unobstructed

visualisation of the entire mucosa. The efficacy of colonoscopy is directly related to adequacy of bowel cleansing; however, ~25% of routine outpatient cases experience suboptimal cleansing.³⁰ The consequences of poor quality BP include decreased rates of

TABLE 2 Reliability of instruments used for the assessment of BP quality in patients with CD.

Instrument	Interrater reliability, ICC (95% CI)	Interpretation	Intrarater reliability, ICC (95% CI)	Interpretation
BBPS				
Insertion	0.63 (0.43, 0.75)	Substantial	0.69 (0.53, 0.81)	Substantial
Withdrawal	0.69 (0.50, 0.79)	Substantial	0.74 (0.59, 0.82)	Substantial
mBBPS				
Insertion	0.69 (0.49, 0.80)	Substantial	0.73 (0.56, 0.84)	Substantial
Withdrawal	0.72 (0.51, 0.81)	Substantial	0.74 (0.57, 0.82)	Substantial
HCS				
Insertion	0.68 (0.48, 0.79)	Substantial	0.80 (0.65, 0.88)	Substantial
Withdrawal	0.69 (0.52, 0.79)	Substantial	0.79 (0.67, 0.86)	Substantial
FDA BCAS				
Insertion	0.51 (0.32, 0.65)	Moderate	0.76 (0.64, 0.84)	Substantial
Withdrawal	0.62 (0.42, 0.74)	Substantial	0.79 (0.67, 0.85)	Substantial
VAS				
Insertion	0.68 (0.50, 0.79)	Substantial	0.83 (0.72, 0.90)	Almost perfect
Withdrawal	0.73 (0.55, 0.83)	Substantial	0.87 (0.77, 0.93)	Almost perfect

Abbreviations: BBPS, Boston Bowel Preparation Scale; BP, bowel preparation; CI, confidence interval; FDA BCAS, Food and Drug Administration Bowel Cleansing Assessment Scale; HCS, Harefield Cleansing Scale; ICC, intraclass correlation coefficient; mBBPS, modified BBPS; VAS, visual analog scale.

TABLE 3 Correlation between instruments used for the assessment of BP quality and the VAS of BP quality/cleanliness.

Instrument	Correlation coefficient (95% CI)
Insertion	
BBPS	0.62 (0.41, 0.74)
mBBPS	0.66 (0.46, 0.78)
HCS	0.66 (0.46, 0.77)
FDA BCAS	0.58 (0.39, 0.70)
Withdrawal	
BBPS	0.70 (0.50, 0.80)
mBBPS	0.71 (0.51, 0.81)
HCS	0.69 (0.51, 0.80)
FDA BCAS	0.67 (0.47, 0.77)

Abbreviations: BBPS, Boston Bowel Preparation Scale; BP, bowel preparation; CI, confidence interval; FDA BCAS, Food and Drug Administration Bowel Cleansing Assessment Scale; HCS, Harefield Cleansing Scale; mBBP, modified BBPS; VAS, visual analog scale.

adenoma detection and caecal intubation, increased procedural complexity including prolonged duration and heightened risk of bleeding necessitating electrocautery, as well as frequent need for repeat colonoscopy.³¹ In addition to dysplasia detection and colorectal cancer screening, optimal bowel cleansing is also imperative for accurate diagnosis and management of CD. Treating to the target of endoscopic healing in CD is a widely accepted strategy associated with improved long-term outcomes and reduced risk of bowel damage.³² Despite clinical remission, ongoing mucosal inflammation can lead to complications, flares, and surgeries. Therefore, regular endoscopic monitoring and appropriate treatment adjustments are

essential for optimal disease management. The European Society of Gastrointestinal Endoscopy (ESGE) recommends specific verbal or written instructions to patients and clinic staff to improve BP quality, and the use of high or low volume PEG-based BP agents in patients with IBD.⁴ A European Crohn's and Colitis Organisation expert panel included optimal BP as a key quality measure, and emphasised the use of standardised scales in comprehensive endoscopic reports.³³

Ideally, a bowel-cleansing product should exhibit safety, efficacy, tolerability, and high compliance with the preprocedural regimen³⁴; however, no studies have been undertaken to evaluate the efficacy of BP products in patients with CD. There is additionally conflicting evidence on the influence of disease-related factors, including endoscopic activity, on BP quality, or conversely, of poor BP quality on the scoring of endoscopic disease activity.^{12,35,36} This has important implications for clinical trials of new therapies for the treatment of CD. Therefore, based on prior research and recommendations from an international expert panel to standardise assessment of bowel cleansing for clinical trials involving patients with CD,¹⁷ we utilised trained central readers to evaluate the operating characteristics of existing BP instruments in patients with CD.

This is the first study to demonstrate that existing instruments used for the assessment of BP quality in the general population are also reliable and valid for use in patients with CD. Specifically, when centrally assessed, the BBPS, mBBPS, HCS, and FDA BCAS demonstrate substantial reliability on both endoscope insertion and withdrawal (with the exception of FDA BCAS which demonstrated moderate inter-rater reliability on withdrawal). While variations in scoring methodologies exist, these differences did not have a noticeable effect on the overall reliability.

TABLE 4 Correlation between instruments used for the assessment of BP quality and the VAS of BP quality/cleanliness according to intestinal segment on the withdrawal phase of the procedure.

Correlation coefficient (95% CI)				
Intestinal segment	BBPS	mBBPS	HCS	FDA BCAS
Ileum	0.40 (0.27, 0.52)	0.43 (0.28, 0.54)	0.46 (0.32, 0.57)	0.39 (0.24, 0.51)
Left colon	0.53 (0.29, 0.67)	0.56 (0.32, 0.70)	0.57 (0.34, 0.70)	0.51 (0.29, 0.65)
Rectum	0.62 (0.27, 0.77)	0.63 (0.26, 0.77)	0.64 (0.49, 0.73)	0.63 (0.28, 0.76)
Right colon	0.73 (0.58, 0.82)	0.62 (0.47, 0.71)	0.71 (0.59, 0.79)	0.72 (0.56, 0.81)
Transverse colon	0.62 (0.47, 0.72)	0.64 (0.48, 0.74)	0.64 (0.49, 0.73)	0.59 (0.43, 0.70)

Abbreviations: BBPS, Boston Bowel Preparation Scale; BP, bowel preparation; CI, confidence interval; FDA BCAS, Food and Drug Administration Bowel Cleansing Assessment Scale; HCS, Harefield Cleansing Scale; mBBPS, modified BBPS.

TABLE 5 Correlation between BP quality and SES-CD scores according to scale and location.

Spearman's rank correlation coefficient (95% CI)					
Scale	Ileum	Right colon	Transverse colon	Left colon	Rectum
BBPS	0.12 (-0.10, 0.33)	-0.16 (-0.39, 0.06)	-0.29 (-0.48, -0.09)	-0.31 (-0.49, -0.13)	-0.14 (-0.33, 0.05)
mBBPS	0.08 (-0.14, 0.30)	-0.10 (-0.33, 0.12)	-0.26 (-0.45, -0.06)	-0.32 (-0.50, -0.14)	-0.13 (-0.33, 0.06)
HCS	0.25 (0.05, 0.45)	-0.12 (-0.34, 0.10)	-0.16 (-0.35, 0.03)	-0.18 (-0.38, 0.02)	-0.05 (-0.27, 0.16)
FDA BCAS	0.11 (-0.12, 0.33)	-0.19 (-0.39, 0.02)	-0.26 (-0.44, -0.08)	-0.33 (-0.51, -0.15)	-0.09 (-0.27, 0.09)

Abbreviations: BBPS, Boston Bowel Preparation Scale; BP, bowel preparation; CI, confidence interval; FDA BCAS, Food and Drug Administration Bowel Cleansing Assessment Scale; HCS, Harefield Cleansing Scale; ICC, intraclass correlation coefficient; mBBPS, modified BBPS; SES-CD, Simple Endoscopic Score for Crohn's Disease.

We also specifically explored the relationship between endoscopic activity and BP quality. We observed negative correlations between BP and endoscopic disease activity scores in the colon. In contrast, BP and endoscopic disease activity were positively correlated in the ileum. A notable negative correlation was observed in the left colon. This finding implies that suboptimal BP in the left colon may lead to overestimation of endoscopic disease activity and underscores the importance of meticulous cleaning in the face of inadequate BP in this segment to ensure an accurate assessment of endoscopic CD activity.¹⁰ This finding contrasts with data reported from a recent large retrospective cohort study demonstrating an independent association between moderate-to-severe IBD disease activity and suboptimal BP (adjusted OR 2.7 [95% CI 1.52–4.94]) relative to mild or no endoscopic disease activity.

Our findings have implications for both clinical trials and clinical practice. Reliable and valid instruments are crucial for accurate evaluation of BP quality in both settings for patients with CD, a population with a unique set of disease-related challenges such as mucosal inflammation, strictures, pseudopolyps, and bowel resection.

Endoscopic evaluation of disease activity is now a component of co-primary endpoints in CD clinical trials. Reliable and valid instruments are therefore essential to ensure accurate assessment and reproducible outcomes.³⁷ Our results provide support for the use of these instruments in clinical trials of new therapies for the treatment of CD, and for the inclusion of patients with CD in trials evaluating new BP formulations. Patients with IBD report decreased tolerability of BP compared with non-IBD patients.⁸ The inclusion of patients

with CD in trials of new BP formulations will ensure that these agents are tolerable, safe, and effective in this unique patient population.

With regard to clinical practice, our findings support the use of existing BP instruments in patients with CD undergoing colonoscopy. Indeed, the overestimation of disease activity due to poor BP raises concerns for clinical decision-making. In post-surgical endoscopy and surveillance, inaccurate disease assessment due to poor BP may lead to overly aggressive treatment intervention with attendant risks that outweigh the potential benefits. Clinicians should be aware of these implications and explore strategies for improving BP quality to ensure appropriate clinical management for patients. Importantly, although we determined that existing BP instruments are all sufficiently reliable for use in patients with CD, the BBPS scale was explicitly designed for use in clinical practice and is the instrument most frequently employed in that setting for determination of BP adequacy, and subsequent need for repeat procedures and/or intervals for screening and surveillance. In contrast, the FDA-BCAS was developed fundamentally for use in clinical trials to assess the efficacy/laxative effect of BP formulations. Although the HCS is theoretically appropriate for either clinical practice or research, given that endoscopists are generally most familiar with the BBPS, and that this instrument is recommended as the preferred instrument in current guidelines,^{5,38} we also recommend the BBPS for BP assessment in patients with CD. This recommendation is additionally based on the prescribed use of the BBPS, which includes assessment after cleaning; an approach that would also be appropriate for assessment of disease activity.

in patients with CD, particularly given the relationship observed between BP quality and endoscopic disease activity scores in this study. Use of the most familiar instrument and/or according to local practice is therefore recommended.

Our study has several strengths. This is the first study to evaluate the operating properties of existing BP instruments in patients with CD, and our findings provide a valuable framework for scoring BP quality in CD clinical trials. The study design was meticulously conceived based on prior research, and incorporated methodology based on expert recommendations.¹⁷ Notably, the results of this study validate the prior recommendations and affirm their relevance in the context of CD. The study population was diverse and included patients with a wide range of disease activity, complications, and prior surgical history. Various procedures were included in our assessment, providing a comprehensive understanding of BP quality across different clinical scenarios. We utilised existing instruments that are well-established, recommended in current guidelines, and validated in clinical practice. Finally, while our results affirm the reliability of BP assessment on endoscope insertion with the FDA BCAS, these results also suggest that reliability of BP assessment is essentially equivalent when assessed on endoscope withdrawal and with any of the other instruments evaluated in this study.

While our study provides valuable insights, we acknowledge some potential limitations. The sample size was relatively small ($N=40$) and included multiple videos from nine patients (eight patients contributed two videos and one patient contributed three videos). However, a total of 50 videos were assessed by the central readers and multiple videos from the same patient were treated as independent observations due to the timing of assessments, use of BP agents, procedure types, and evolution of disease-related factors between videos. Furthermore, a sensitivity analysis that included a single video from each patient found no difference in study outcomes (data not shown).

Furthermore, the expertise and training of the central readers in this study may have resulted in reliability estimates that may differ from real-world scenarios. Also, while our study explores assumptions based on the correlation between BP quality and disease activity, it must be acknowledged that these associations are exploratory and founded on small correlations.

Future research might explore the impact of specific CD-related factors (e.g., the presence of strictures) on BP quality and involve a larger and more diverse sample to enhance the generalisability of the findings. Further investigations might also assess the effectiveness of interventions aimed at optimising BP in patients with CD, including patient education.

In conclusion, this study contributes important new information on the operating properties of existing BP quality instruments in patients with CD. These results, and those on the potential impact of BP quality on the quantification of endoscopic disease activity in patients with CD have important implications for both clinical practice and clinical research. Our findings, which should help to foster the inclusion of CD patients in clinical trials of new BP formulations and facilitate standardised assessment of BP quality in clinical trials of new therapies, are significant steps towards improving the quality of care of patients with CD.

AUTHOR CONTRIBUTIONS

Virginia Solitano: Conceptualization; writing – original draft; formal analysis; writing – review and editing; methodology. **Corey A. Siegel:** Conceptualization; formal analysis; writing – review and editing; methodology. **Joshua R. Korzenik:** Conceptualization; formal analysis; writing – review and editing; methodology. **Jennifer K. Maratt:** Writing – review and editing. **Douglas K. Rex:** Writing – review and editing. **Bryan Maguire:** Formal analysis; writing – review and editing. **Brian Bressler:** Writing – review and editing. **Johannes Grossmann:** Writing – review and editing. **Rocio Sedano:** Writing – review and editing. **John W. D. McDonald:** Writing – review and editing; project administration. **Julie Remillard:** Writing – review and editing; project administration. **Lisa M. Shackelton:** Writing – review and editing; formal analysis; writing – original draft. **Guangyong Zou:** Formal analysis; methodology. **Brian G. Feagan:** Conceptualization; writing – review and editing; methodology. **Christopher Ma:** Writing – review and editing; methodology. **Vipul Jairath:** Conceptualization; writing – review and editing; formal analysis; writing – original draft; supervision; funding acquisition; methodology.

ACKNOWLEDGEMENTS

This project was supported in part by the Leona M. and Harry B. Helmsley Charitable Trust (Prime award no. #2005-03895).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

AUTHORSHIP

Guarantor of the article: Dr. Vipul Jairath.

ORCID

Jennifer K. Maratt  <https://orcid.org/0000-0001-9218-7391>

Christopher Ma  <https://orcid.org/0000-0002-4698-9948>

Vipul Jairath  <https://orcid.org/0000-0002-1092-0033>

REFERENCES

- Lebwohl B, Kastrinos F, Glick M, Rosenbaum AJ, Wang T, Neugut AI. The impact of suboptimal bowel preparation on adenoma miss rates and the factors associated with early repeat colonoscopy. *Gastrointest Endosc.* 2011;73:1207–14.
- Harewood GC, Sharma VK, de Garmo P. Impact of colonoscopy preparation quality on detection of suspected colonic neoplasia. *Gastrointest Endosc.* 2003;58:76–9.
- Sulz MC, Kroger A, Prakash M, Manser CN, Heinrich H, Misselwitz B. Meta-analysis of the effect of bowel preparation on adenoma detection: early adenomas affected stronger than advanced adenomas. *PLoS One.* 2016;11:e0154149.
- Hassan C, East J, Radaelli F, Spada C, Benamouzig R, Bisschops R, et al. Bowel preparation for colonoscopy: European Society of Gastrointestinal Endoscopy (ESGE) guideline—update 2019. *Endoscopy.* 2019;51:775–94.
- Keswani RN, Crockett SD, Calderwood AH. AGA clinical practice update on strategies to improve quality of screening and surveillance colonoscopy: expert review. *Gastroenterology.* 2021;161:701–11.

6. Negreanu L, Voiosu T, State M, Voiosu A, Bengus A, Mateescu BR. Endoscopy in inflammatory bowel disease: from guidelines to real life. *Ther Adv Gastroenterol*. 2019;12:1756284819865153.
7. Solitano V, Zilli A, Franchellucci G, Allocacca M, Fiorino G, Furfaro F, et al. Artificial endoscopy and inflammatory bowel disease: welcome to the future. *J Clin Med*. 2022;11:569.
8. Mohsen W, Williams AJ, Wark G, Sechi A, Koo JH, Xuan W, et al. Prospective single-blinded single-center randomized controlled trial of prep kit-C and Moviprep: does underlying inflammatory bowel disease impact tolerability and efficacy? *World J Gastroenterol*. 2021;27:1090–100.
9. Denters MJ, Schreuder M, Depla AC, Mallant-Hent RC, van Kouwen MC, Deutekom M, et al. Patients' perception of colonoscopy: patients with inflammatory bowel disease and irritable bowel syndrome experience the largest burden. *Eur J Gastroenterol Hepatol*. 2013;25:964–72.
10. Kumar A, Shenoy V, Buckley MC, Durbin L, Mackey J, Mone A, et al. Endoscopic disease activity and biologic therapy are independent predictors of suboptimal bowel preparation in patients with inflammatory bowel disease undergoing colonoscopy. *Dig Dis Sci*. 2022;67:4851–65.
11. Khanna R, Zou G, D'Haens G, Rutgeerts P, McDonald JW, Daperno M, et al. Reliability among central readers in the evaluation of endoscopic findings from patients with Crohn's disease. *Gut*. 2016;65:1119–25.
12. Manes G, Fontana P, de Nucci G, Radaelli F, Hassan C, Ardizzone S. Colon cleansing for colonoscopy in patients with ulcerative colitis: efficacy and acceptability of a 2-L PEG plus Bisacodyl versus 4-L PEG. *Inflamm Bowel Dis*. 2015;21:2137–44.
13. Lazzaroni M, Petrillo M, Desideri S, Bianchi Porro G. Efficacy and tolerability of polyethylene glycol-electrolyte lavage solution with and without simethicone in the preparation of patients with inflammatory bowel disease for colonoscopy. *Aliment Pharmacol Ther*. 1993;7:655–9.
14. Kastenberg D, Bertiger G, Brogadir S. Bowel preparation quality scales for colonoscopy. *World J Gastroenterol*. 2018;24:2833–43.
15. Hanzel J, Sey M, Ma C, Zou G, East JE, Siegel CA, et al. Existing bowel preparation quality scales are reliable in the setting of centralized endoscopy Reading. *Dig Dis Sci*. 2023;68:1195–207.
16. U.S. Food and Drug Administration. Bowel cleansing for colonoscopy: efficacy and safety considerations for developing new products. 2021.
17. Maratt JK, Siegel CA, Barkun AN, Bouhnik Y, Bressler B, Calderwood AH, et al. An expert consensus to standardize assessment of bowel cleansing for clinical trials of bowel preparations for Crohn's disease. *Dig Dis Sci*. 2023;68:1718–27.
18. 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.
19. Calderwood AH, Jacobson BC. Comprehensive validation of the Boston Bowel Preparation Scale. *Gastrointest Endosc*. 2010;72:686–92.
20. Halphen M, Heresbach D, Gruss HJ, Belsey J. Validation of the Harefield Cleansing Scale: a tool for the evaluation of bowel cleansing quality in both research and clinical practice. *Gastrointest Endosc*. 2013;78:121–31.
21. Daperno M, D'Haens G, Van Assche G, Baert F, Bulois P, Maunoury V, et al. Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. *Gastrointest Endosc*. 2004;60:505–12.
22. Gilder KTN, Tian LL, Cappelleri JC, Hanumara RC. Confidence intervals on intraclass correlation coefficients in a balanced two-factor random design. *J Stat Plan Infer*. 2007;137:1199–212.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–74.
24. Kowalik DCY-H, Zou G. Confidence interval estimation for a difference between two dependent Intraclass correlation coefficients with variable class sizes. *J Stat Theory Pract*. 2010;5:613–25.
25. Huang Q. PROC MIXED: calculate correlation coefficients in the presence of repeated measurements. 2019.
26. Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: part 2—correlation between subjects. *BMJ*. 1995;310:633.
27. Zou GY. Toward using confidence intervals to compare correlations. *Psychol Methods*. 2007;12:399–413.
28. Cohen J. A power primer. *Psychol Bull*. 1992;112:155–9.
29. Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. *Stat Med*. 2012;31:3972–81.
30. van Riswijk MLM, van Keulen KE, Siersema PD. Efficacy of ultra-low volume (</=1 L) bowel preparation fluids: systematic review and meta-analysis. *Dig Endosc*. 2022;34:13–32.
31. Johnson DA, Barkun AN, Cohen LB, Dominitz JA, Kaltenbach T, Martel M, et al. Optimizing adequacy of bowel cleansing for colonoscopy: recommendations from the US multi-society task force on colorectal cancer. *Gastroenterology*. 2014;147:903–24.
32. Turner D, Ricciuti A, Lewis A, D'Amico F, Dhaliwal J, Griffiths AM, et al. STRIDE-II: an update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) initiative of the International Organization for the Study of IBD (IOIBD): determining therapeutic goals for treat-to-target strategies in IBD. *Gastroenterology*. 2021;160:1570–83.
33. Adamina M, Feakins R, Iacucci M, Spinelli A, Cannatelli R, D'Hoore A, et al. ECCO topical review Optimising reporting in surgery, endoscopy, and histopathology. *J Crohns Colitis*. 2021;15:1089–105.
34. Wexner SD, Beck DE, Baron TH, Fanelli RD, Hyman N, Sher B, et al. A consensus document on bowel preparation before colonoscopy: prepared by a task force from the American Society of Colon and Rectal Surgeons (ASCRS), the American Society for Gastrointestinal Endoscopy (ASGE), and the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES). *Gastrointest Endosc*. 2006;63:894–909.
35. Negreanu L, Voiosu T, State M, Mateescu RB. Quality of colonoscopy preparation in patients with inflammatory bowel disease: retrospective analysis of 348 colonoscopies. *J Int Med Res*. 2020;48:300060520903654.
36. Bessisow T, Van Keerberghen CA, Van Oudenhove L, Ferrante M, Vermeire S, Rutgeerts P, et al. Anxiety is associated with impaired tolerance of colonoscopy preparation in inflammatory bowel disease and controls. *J Crohns Colitis*. 2013;7:e580–e587.
37. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gotzsche PC, Krle AJK, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Rev Panam Salud Publica*. 2015;38:506–14.
38. Kaminski MF, Thomas-Gibson S, Bugajski M, Brethauer M, Rees CJ, Dekker E, et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) quality improvement initiative. *United European Gastroenterol J*. 2017;5:309–34.

SUPPORTING INFORMATION

Additional supporting information will be found online in the Supporting Information section.

How to cite this article: Solitano V, Siegel CA, Korzenik JR, Maratt JK, Rex DK, Maguire B, et al. Performance of bowel preparation quality scales in patients with Crohn's disease. *Aliment Pharmacol Ther*. 2024;60:1042–1050. <https://doi.org/10.1111/apt.18210>