

KODA score: an updated and validated bowel preparation scale for patients undergoing small bowel capsule endoscopy




Authors

Mohammad Alageeli¹, Brian Yan¹, Suliman Alshankiti², May Al-Zahrani¹, Zoya Bahreini¹, ThucNhi T. Dang, Joshua Friedland³, Syed Gilani³, Ravi Homenauth², Jennifer Houle⁴, Milica Kloc⁴, Jacqueline Luhoway³, Lucas Merotto³, Rymon Rofaie⁵, Chantal Singh⁴, Alexis Smith⁶, Benson Thomas¹, Cassandra Townsend¹, David Yoo⁵, Sergio Zepeda-Gomez², Larry Stitt⁷, Vipul Jairath^{1,7,8,9}, Michael Sai Lai Sey^{1,8}

Institutions

- 1 Division of Gastroenterology, Western University, London, Ontario, Canada
- 2 Division of Gastroenterology, University of Alberta, Edmonton, Alberta, Canada
- 3 Schulich School of Medicine, Western University, London, Ontario, Canada
- 4 London Health Sciences Centre, London, Ontario, Canada
- 5 Department of Medicine, London Health Sciences Center, London, Ontario, Canada
- 6 Arthur Labatt School of Nursing, Western University, London, Ontario, Canada
- 7 Robarts Clinical Trials Inc., London, Ontario, Canada
- 8 Program for Experimental Medicine, Western University, London, Ontario, Canada
- 9 Department of Epidemiology and Biostatistics, Western University, London, Ontario, Canada

submitted 13.11.2019

accepted after revision 1.4.2020

Bibliography

DOI <https://doi.org/10.1055/a-1176-9889> |
 Endoscopy International Open 2020; 08: E1011–E1017
 © Georg Thieme Verlag KG Stuttgart · New York
 eISSN 2196-9736

Corresponding author

Michael Sey MD, MPH, FRCP(C), Assistant Professor of Medicine, Western University, Small Intestinal Endoscopy Program Director, Division of Gastroenterology, London Health Sciences Centre-Victoria Hospital, 800 Commissioners Rd. E., London, ON, Canada, N6A 5W9
 Fax: +519-667-6820
msey2@uwo.ca

ABSTRACT

Background and study aims A reliable outcome measure is needed for bowel preparation quality during capsule endoscopy. Currently, no scales are adequately validated. Our objective was to update an existing small bowel preparation score, create a standardized training module, then determine its inter-rater and intra-rater reliability.

Patients and methods Modification to produce standardized scoring of an existing small bowel preparation score was performed followed by development of a training module and validation to create the new Korea-Canada (KODA) score. Twenty readers from a range of backgrounds, including capsule endoscopists, gastroenterology fellows, residents, medical students, and nurses rated bowel cleanliness in 25 capsule videos consisting of 1,233 images, in duplicate 4 weeks apart, after completing the training module. Sequential images selected in 5-minute intervals during small bowel transit were rated on a scale between 0–3 based on the amount of visualized mucosa and the degree of obstruction. Reliability was assessed using estimates of intraclass correlation coefficients (ICCs).

Results Intraclass correlation coefficients for inter-rater (ICC 0.81, 95% CI 0.70–0.87) and intra-rater (ICC 0.92, 95% CI 0.87–0.94) reliability were almost perfect among the 20 readers. Inter-rater reliability ranged between 0.72 (95% CI 0.57–0.81) and 0.89 (95% CI 0.79–0.93) for nurses and residents, respectively. Intra-rater reliability was greater than 0.90 for all groups except for nurses, which was still almost perfect (ICC 0.86, 95% CI 0.79–0.90).

Conclusions Almost perfect inter-rater and intra-rater reliability was observed for the KODA score. This simple score could be used for future clinical trials after completion of the training module.

Introduction

The advent of capsule endoscopy nearly two decades ago revolutionized the investigation of small bowel diseases [1, 2]. Similar to colonoscopy, adequate bowel preparation is essential for a quality exam, albeit in the small intestine rather than the colon [3]. However, unlike colonoscopy in which numerous bowel preparation scores exist, there is no validated or widely accepted scale for bowel preparation in the small intestine [4]. As a result, research in bowel preparation for capsule endoscopy, including clinical trials, have utilized a range of ad hoc qualitative and semi-quantitative scales [5–8]. This is problematic as the reliability and operating properties of such scales are unknown. A reliable small bowel preparation scale should produce a similar score if measured by different outcome assessors (inter-rater reliability) or by the same assessor at different times (intra-rater reliability) [9]. Furthermore, use of numerous scales in the literature has resulted in substantial heterogeneity in meta-analyses of studies investigating small bowel preparation, resulting in inconsistent conclusions [10, 11]. Accordingly, a reliable and valid measure of small bowel preparation would improve the quality of capsule endoscopy research.

Small bowel preparation assessment is difficult due to the long duration of capsule endoscopy videos, where small intestinal transit often take hours to complete [11]. Accordingly, small bowel preparation scales that require review of the entire video for scoring, such as using a stopwatch to record the proportion of time the mucosa is obscured by intestinal contents [6, 12, 13], are unsuitable for use due to being overly labor intensive. As a simplification, some scores rate small bowel cleanliness using an ordinal scale, usually some variation of excellent, good, fair, and poor, for the entire small bowel based on an overall impression or “gestalt” [7, 14, 15]. However, such a simple scale is likely insufficiently discriminative for use. This has been shown in the colonoscopy literature where bowel preparation scales that require assessment of individual segments of the colon, such as the Boston Bowel Preparation Scale or the Ottawa Bowel Preparation Quality Scale, were shown to have interobserver reliability superior to the Aronchick Scale, which relies only on overall assessment of the colon as a whole [4]. The advantage of segmental compared to overall assessment could be even greater in the small intestine due to its length compared to the relatively short colon. Other strategies employed have taken advantage of segmental scoring, such as dividing the video into five parts and only scoring 5-minute segments within each part [5]; dividing the video into quartiles by time and evaluating the first and last 10-minute segment in each part [16]; evaluating only the first and last hour of small intestinal transit [17]; dividing the video into 10 parts by time and scoring the first 5 minutes of each segment [18].

The most promising small bowel preparation scale to date was developed by Park and colleagues [19]. Unlike other scales, theirs requires selection of consecutive images at 5-minute intervals, thereby sampling the entire small intestine in many small segments. The images are scored between 0 to 3 based on two domains: proportion of visualized mucosa in the image (VM) and the degree of obstruction by bubbles, debris, and bile

(DO) (► **Fig. 1a**). The final score is a mean of the two domain sub-scores. They reported excellent inter-rater agreement (ICC=0.82) and intra-rater agreement (ICC>0.80) for this scale. Furthermore, to ensure the sampled images are representative of the overall video, they compared it to a very labor-intensive strategy in which every image within the first 2 minutes of every 5-minute segment was assessed (e.g. 240 images/5-minute segment) and reported excellent agreement between the two (ICC 0.82).

Although promising, the operating characteristics of the score have not been validated in an independent cohort made up of different patients and capsule readers. This step is critical to ensuring the operating characteristics of the score, including inter-rater and intra-rater reliability, are generalizable beyond the original investigators. Thus, we aimed to independently validate the reliability of the scale using a diverse range of readers and to develop a training module to permit future dissemination of the score.

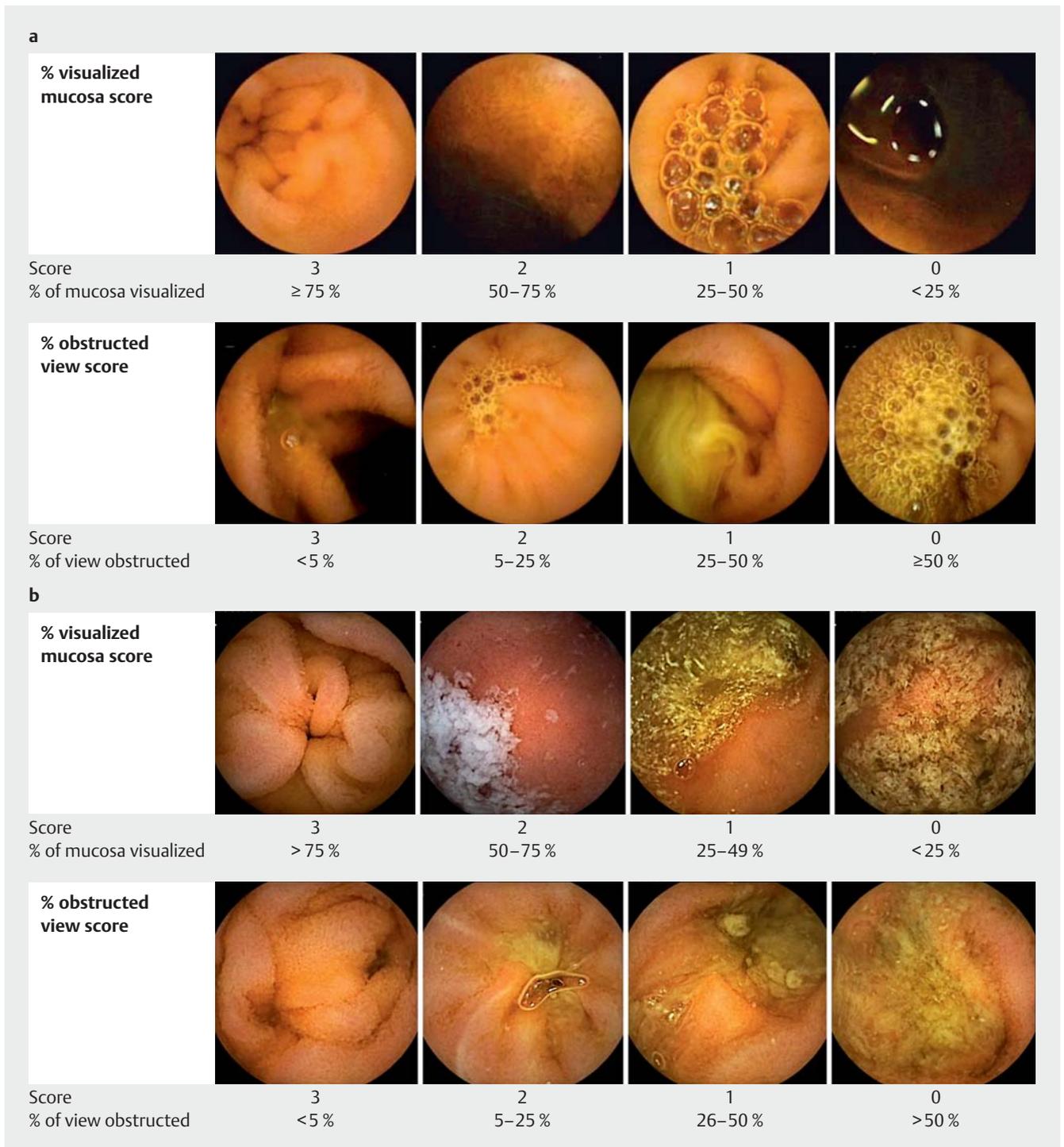
Patients and methods

Study design

The study was conducted between 2016 and 2017 and consisted of three phases: Training Phase, Assessment Module 1, and Assessment Module 2. For the training phase, we created a bespoke online platform which contained instructions on how to rate small bowel cleanliness using the study scale in a standardized manner and five practice videos to score with feedback for the first three (https://www.schulich.uwo.ca/gastroenterology/research/research_tools). One week following completion of the Training Module, readers were sent a link to Assessment Module 1, which consisted of images from 25 capsule videos. Each reader independently rated all images using the study scale and they were unaware of any clinical information related to the capsule video. Four weeks after completion of Assessment Module 1, readers were sent a link to Assessment Module 2, which consisted of the same images as Assessment Module 1 to measure inter-rater and intra-rater reliability. All modules were web-based (www.Qualtrics.com) and could be completed at the time and location of the reader's choosing. The only requirement was to avoid using a cellular phone due to the small screen size. Readers were given 1 week to complete the Training Module and 2 weeks for each Assessment Module given the large number of images involved. The study was approved by the Western University Research Ethics Board (REB# 108350).

Selection of Readers

To maximize generalizability, individuals with varying levels of experience reading capsule endoscopy, at different levels of training, and in different health care professions, were selected as readers. The readers included four capsule endoscopists, who were gastroenterologists with expertise in capsule endoscopy and deep enteroscopy; four gastroenterology fellows; four internal medicine residents; four medical students; and four registered nurses. These groups were selected based on the likelihood they would be involved in capsule endoscopy reading



► **Fig. 1** **a** Original iteration of small bowel preparation score [19]. **b** KODA score.

and research. Readers either held academic appointments, were training, or employed at Western University (London, Canada) or the University of Alberta (Edmonton, Canada) or their affiliated hospitals during the study period. Demographic information was collected for each reader along with the number of capsules they had read before study initiation.

Preparation of capsule videos

A total of 1,233 images from 25 consecutive capsule videos performed at London Health Sciences Centre, a tertiary care hospital affiliated with Western University (London, Canada), between July and October 2015 were used to develop the Assessment Modules. Capsule studies that failed to reach the cecum were excluded. Demographic information regarding the

patients and indications for capsule endoscopy were noted. All patients received bowel preparation consisting of 2 L polyethylene glycol electrolyte solution (PegLyte, Pharmascience Inc., Montreal, Canada) the night before capsule endoscopy and clear fluids until fasting started at midnight. In each video, the first duodenal image and first cecal image were landmarked. Thereafter, the first image at 5-minute intervals during small intestinal transit, defined as the period between the first duodenal image and first cecal image, was selected. In cases where the capsule refluxed back into the stomach, the first duodenal image was selected as the last time the capsule entered the small bowel. If the capsule appeared to remain in one location for more than 5 minutes, only one image was selected rather than repeatedly capturing the same image. Images were de-identified and exported as high-quality .jpg files before being uploaded into the Assessment Modules. All capsule studies were performed using the PillCamSB3 capsule and read on RAPID 8.0 (Medtronic, Minneapolis, Minnesota, United States).

Small bowel preparation scale

Each individual image was scored on two domains: visualized mucosa (VM), defined as the percentage of mucosa visible in the image, and degree of obstruction (DO), defined as the percentage of the image obscured by debris, bubbles, and bile. Each domain was assigned a score between 0 and 3, and the overall score was the mean of the two domain scores (► **Fig. 1b**). Two minor modifications were made from the original description [19] and defined a priori (► **Fig. 1a**). First, to avoid possible ambiguity, scoring was modified to ensure that levels were mutually exclusive. For example, in the original iteration, VM was scored as 3 if $\geq 75\%$ of the mucosa could be visualized and 2 if 50% to 75% of the mucosa was visualized, permitting a score of 3 or 2 to be assigned for an image with exactly 75% of the mucosa visualizable. As such, we adjusted the cut-off values to maintain mutual exclusivity (ex. 3 = $> 75\%$ of the mucosa visualized, 2 = 50% to 75% of mucosa visualized). Second, we formalized the handling of shadows, which was not clear in the original publication in terms of whether it affects the VM sub-score, DO sub-score, or both. Sample images from the original iteration of the score seem to suggest that shadows should only penalize the VM sub-score and not the DO sub-score (ex. ► **Fig. 1a**, DO score of 3 was rated as having $< 5\%$ of view obstructed despite there being a large shadow present). Intuitively, we felt this made sense since a shadow would reduce the amount of mucosa visualized but not obstruct the view per se. As such, we amended the scoring guidelines to explicitly state that the presence of shadows should affect the VM sub-score but not the DO sub-score.

Objectives

The primary objective of the study was to estimate the inter-rater and intra-rater reliability of the study scale. Secondary objectives included estimation of the inter-rater and intra-rater reliabilities within each reader group and the median (IQR) for the overall score and sub-scores for all readers and by reader group.

Statistical analyses

Inter-rater and intra-rater reliability estimates measured by the intraclass correlation coefficients (ICC) and their corresponding 95% confidence intervals for the total study score, and for the VM and DO sub-scores, were estimated for capsule endoscopists, gastroenterology fellows, internal medicine residents, medical students, and registered nurses separately and combined. Point estimates were obtained using a two-way random effects model with interaction between videos and readers as described by Eliasziw et al [20]. To avoid the normality assumption, the associated two-sided 95% confidence intervals for the reliability coefficients were obtained using the non-parametric percentile bootstrap method [21], commonly known as the cluster bootstrap method, with 2000 replicates, sampled and replaced at the level of the video.

The strength of the reliability estimates was interpreted according to the subjective but well-established, benchmarks of Landis and Koch [22] whereby ICC of < 0.00 , 0.00–0.20, 0.21–0.40, 0.41–0.60, 0.61–0.80, above 0.80 indicate poor, slight, fair, moderate, substantial and almost perfect reliability, respectively.

Sample size for reliability was based on the one-way random effects model, which tends to provide more liberal estimates compared to those based on the two-way models. Assuming a true ICC of 0.8, evaluation of 25 videos 2 times by 20 readers would yield an approximately 90% chance of obtaining a lower bound for the two-sided 95% CI for an ICC greater than 0.6 [23].

Results

Reader demographics

The median (range) age of readers was 30 (22–49) and 40% were females. Capsule endoscopists read a median (range) of 205 capsules (50–500) before study initiation. Among gastroenterology fellows, the median (range) number of capsules read was 6 [5–7]. Internal medicine residents, medical students, and registered nurses reported no capsule reading experience at the start of the study. All 20 readers completed the three phases of the study. The mean (SD) time required to complete the training module was 65.2 minutes (20.2) after exclusion of a single outlier. In this case, the outlying reader was confirmed later to have left his web browser open while not working on the module, thus inadvertently logging a total of 7225 minutes on the training module.

Patient demographics

The median age (range) of capsule endoscopy patients in the study was 67.5 (26, 85) and 14 (56%) were females. Fifteen patients (60%) had occult obscure gastrointestinal bleeding (OGIB), six (25%) had overt OGIB, two (8%) had Crohn's disease, and two (8%) underwent small bowel screening for a hereditary polyposis syndrome. The median (range) gastric emptying time was 18 minutes (1, 140) and small bowel transit time 220 minutes (33, 445), respectively.

► **Table 1** Median (IQR) scores assigned by the study cohort.

	Median (IQR)		
	Visualized mucosa sub-score	Degree of obstruction sub-score	Overall score
Capsule endoscopist	1.49 (1.20, 1.79)	1.83 (1.51, 2.25)	1.69 (1.35, 2.01)
Gastroenterology fellows	1.52 (1.27, 1.86)	1.91 (1.57, 2.38)	1.71 (1.44, 2.11)
Internal medicine residents	1.74 (1.41, 2.01)	2.18 (1.89, 2.59)	1.96 (1.68, 2.27)
Medical students	1.56 (1.28, 1.88)	1.87 (1.52, 2.33)	1.73 (1.42, 2.12)
Registered nurses	1.60 (1.29, 1.88)	1.99 (1.64, 2.45)	1.81 (1.48, 2.17)
Overall	1.58 (1.28, 1.89)	1.96 (1.61, 2.41)	1.77 (1.46, 2.14)

IQR, interquartile range

► **Table 2** Reliability estimates for the KODA score.

	ICC (95% CI)					
	Visualized mucosa sub-score		Degree of obstruction sub-score		Overall score	
	Inter-rater	Intra-rater	Inter-rater	Intra-rater	Inter-rater	Intra-rater
Capsule endoscopists	0.65 (0.50, 0.75)	0.93 (0.89, 0.96)	0.90 (0.83, 0.93)	0.94 (0.89, 0.96)	0.85 (0.76, 0.90)	0.93 (0.89, 0.96)
Gastroenterology Fellows	0.86 (0.79, 0.90)	0.89 (0.83, 0.93)	0.81 (0.69, 0.88)	0.91 (0.87, 0.94)	0.87 (0.80, 0.92)	0.92 (0.87, 0.95)
Internal Medicine residents	0.83 (0.72, 0.89)	0.92 (0.87, 0.95)	0.86 (0.75, 0.91)	0.95 (0.92, 0.96)	0.89 (0.79, 0.93)	0.95 (0.92, 0.97)
Medical students	0.89 (0.80, 0.93)	0.95 (0.91, 0.97)	0.79 (0.68, 0.86)	0.93 (0.88, 0.95)	0.88 (0.80, 0.93)	0.94 (0.90, 0.96)
Registered nurses	0.76 (0.61, 0.83)	0.85 (0.77, 0.90)	0.67 (0.51, 0.77)	0.85 (0.77, 0.90)	0.72 (0.57, 0.81)	0.86 (0.79, 0.90)
Overall	0.79 (0.67, 0.85)	0.91 (0.86, 0.93)	0.77, (0.64, 0.84)	0.91 (0.87, 0.94)	0.81 (0.70, 0.87)	0.92 (0.87, 0.94)

ICC, intraclass correlation coefficient; CI, confidence interval

Bowel preparation score and reliability estimates

The median (IQR) score for small bowel cleanliness was 1.77 (1.46, 2.14) for the 20 readers and 1.69 (1.35, 2.01), 1.71 (1.44, 2.11), 1.96 (1.68, 2.27), 1.73 (1.42, 2.12), 1.81 (1.48, 2.17) for capsule endoscopists, gastroenterology fellows, internal medicine residents, medical students, and RNs, respectively (► **Table 1**). The median (IQR) VM sub-score was 1.58 (1.28, 1.89) and the median (IQR) DO sub-score was 1.96 (1.61, 2.41).

For the readers overall, the inter-rater reliability (ICC 0.81, 95%CI 0.70–0.87) and intra-rater reliability (ICC 0.92, 95% CI 0.87–0.94) of the study scale was almost perfect (► **Table 2**). For the VM sub-score, the inter-rater reliability (ICC 0.79, 95% CI 0.67–0.85) was substantial and the intra-rater reliability (ICC 0.91, 95% CI 0.86–0.93) was almost perfect. Similarly, for the DO sub-score, the inter-rater reliability (ICC 0.77, 95% CI 0.64–0.84) was substantial and the intra-rater reliability (ICC 0.91, 95% CI 0.87–0.94) was almost perfect.

Discussion

We performed the first validation of the reliability of a small bowel preparation scale. To reflect the work by the original Korean investigators who developed the score and its subsequent

validation and training module development in Canada, we propose, in conjunction with our Korean colleagues (e-mail communication with corresponding author, Dr. Bora Keum), naming this scale the KOREA-CANADA (KODA) score.

A reliable outcome measure should be used in research studies, especially clinical trials. Unvalidated outcome measures with unknown inter-rater and intra-rater reliabilities may yield vastly different results if measured by different outcome assessors or even by the same assessor at different times. The lack of consistency in how an outcome would be rated in these scenarios seriously threatens the validity of findings from such studies. To our knowledge, there have only been two prior attempts at validating a small bowel preparation scale. The first was by Hong-Bin et al. [24], who reported on the reliability of the Visualized Area Percentage Assessment of Cleansing Score (AAC) and Computed Assessment of Cleansing (CAC) score. However, both scores require specialized image processing software beyond the standard capsule endoscopy reading program and the reliability estimates were based on only two readers. The second was a validation study of a 5-point ordinal small bowel cleanliness scale performed by three readers as part of a clinical trial protocol [25]. To date, the study has only been published in abstract form, thus not allowing adequate assessment of its operating characteristics.

In comparison, our study validated the reliability of the KODA score among 20 readers of varying backgrounds reviewing 1,233 images from 25 capsule videos twice and reported almost perfect inter-rater and intra-rater reliability. The inter-rater and intra-rater reliabilities were almost perfect in individual reader groups with the exception of nurses, which was nonetheless substantial. In clinical practice, capsule endoscopists, gastroenterology fellows, and nurses are the most likely groups to review capsules. However, we included internal medicine residents and medical students for two reasons. First, it is conceivable they may be involved in research that would require them to use the KODA score. Second, and perhaps more importantly, the fact that the score performed so well in these individuals with no endoscopy or gastroenterology experience supports the generalizability of it overall to diverse readers of varied backgrounds.

The primary strength of the KODA score beyond reliability is simplicity and ease of use. Prior attempts at measuring small bowel cleanliness either relied on manual review of the entire capsule video or required specialized imaging software beyond that routinely available in clinical practice [7, 15, 24, 26–34]. Given that the former may require in excess of 1 hour to review a single capsule video for bowel preparation quality and the latter requires extraction of high-quality screen captures for further image processing using specialized software more commonly used by graphic designers, neither are feasible options for most clinicians. In contrast, the KODA score only requires review of a single image for every 5 minutes of small bowel transit. As such, in a typical capsule video with 3 hours of small bowel transit, only 36 images need to be reviewed. Image selection is made rapid by typing in 5-minute intervals in the time field of the video player after the first duodenal image, negating the need to manually scroll through the video to select images. In our experience, a video typically requires less than 60 seconds to mark images for scoring. In addition, everything can be accomplished within the native capsule endoscopy computer software already in use for clinical practice. Overall, the KODA score is sufficiently easy to use that even readers with no capsule experience, including students, were able to achieve substantial to almost perfect inter-rater and intra-rater reliability after completing the training module. Thus, our results indicate that the KODA score can be used by a diverse range of people upon completion of the Training Module we developed.

There are several limitations of the KODA score and our study. First, the KODA score may be criticized for not reviewing the video in its entirety and instead only rating sequential images 5 minutes apart. Overall, we view this as a benefit of the score given that it saves substantial time. Perhaps more importantly, the representativeness of reviewing images every 5 minutes to the overall cleanliness of the small intestine has already been established. Park et al. [19] previously demonstrated this by reporting almost perfect agreement (ICC 0.82) between scoring images every 5 minutes and scoring every image within the first 2 minutes of every 5-minute segment (e.g. 11,520 images scored in a standard 4-hour video). Ultimately, given the long duration of capsule studies, most scoring systems employ some mechanism to permit scoring of bowel preparation

quality without the need to review the video in its entirety, such as dividing the video into five parts and only scoring 5-minute segments within each part [5]; dividing the video into quartiles and evaluating the first and last 10-minute segment in each part [16]; evaluating only the first and last hour of small intestinal transit [17]; dividing the video into 10 parts and scoring the first 5 minute of each segment [18], or using a colorimetric assay of the progress bar without reviewing the images themselves [24]. Given the correlation between scoring images every 5 minutes and the overall small bowel preparation quality for our score, which has already been established with almost perfect agreement, we do not feel this aspect of our score threatens its validity.

Second, although we demonstrated substantial reliability of the KODA score, we did not address other aspects of validity. Given that the score measures the two most important aspects of capsule imaging quality, the proportion of mucosa visualized and the degree of view that is obstructed, we feel that the KODA score has at least face validity. However, further work will be required to correlate the score with clinical outcomes, which is beyond the scope of this reliability study. Third, we were unable to randomize the order of the capsule studies reviewed during Assessment Module 2 owing to the large number of images involved in the study. However, given the 4-week interval between Assessment Modules 1 and 2, and the fact that 1,233 images were read in duplicate, recall bias during Assessment Module 2 would be very difficult and highly unlikely.

Conclusion

In conclusion, the KODA score is an easy-to-use and highly reliable scale for assessing small bowel preparation quality during capsule endoscopy. Readers of varied background can achieve competency in using this scale after completion of the training module (assessable free-of-charge at https://www.schulich.uwo.ca/gastroenterology/research/research_tools). We recommend its use in future clinical trials for capsule endoscopy to enable standardization of bowel preparation quality assessment.

Acknowledgements

The authors thank Dr. Bora Keum and colleagues who initiated the work on the KoDa score. The study was supported by a Health Research Grant from the PSI Foundation.

Competing interests

Michael Sai Lai Sey – Received arms-length research grant from Pharmascience Inc. and educational grants from Cook Medical, Medtronic, and Olympus. All other authors declare that they have no conflict of interest.

References

- [1] Iddan G, Meron G, Glukhovskiy A et al. Wireless capsule endoscopy. *Nature* 2000; 405: 417
- [2] Pennazio M, Spada C, Eliakim R et al. Small-bowel capsule endoscopy and device-assisted enteroscopy for diagnosis and treatment of small-bowel disorders: European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline. *Endoscopy* 2015; 47: 352–376
- [3] Enns RA, Hookey L, Armstrong D et al. Clinical Practice Guidelines for the Use of Video Capsule Endoscopy. *Gastroenterology* 2017; 152: 497–514
- [4] Parmar R, Martel M, Rostom A et al. Validated scales for colon cleansing: a systematic review. *Am J Gastroenterol* 2016; 111: 197–204
- [5] Lapalus MG, Ben Soussan E, Saurin JC et al. Capsule endoscopy and bowel preparation with oral sodium phosphate: a prospective randomized controlled trial. *Gastrointest Endosc* 2008; 67: 1091–1096
- [6] Wei W, Ge ZZ, Lu H et al. Purgative bowel cleansing combined with simethicone improves capsule endoscopy imaging. *Am J Gastroenterol* 2008; 103: 77–82
- [7] Pons Beltran V, Gonzalez Suarez B, Gonzalez Asanza C et al. Evaluation of different bowel preparations for small bowel capsule endoscopy: a prospective, randomized, controlled study. *Dig Dis Sci* 2011; 56: 2900–2905
- [8] Chen HB, Lian-Xiang P, Yue H et al. Randomized controlled trial of 3 days fasting and oral senna, combined with mannitol and simethicone, before capsule endoscopy. *Medicine* 2017; 96: e8322
- [9] Paul C, Price RSJ, Chiang IA. *Research Methods in Psychology*. 2nd Canadian Edition. The Saylor Foundation; 2015: 90–94
- [10] Kotwal VS, Attar BM, Gupta S et al. Should bowel preparation, anti-foaming agents, or prokinetics be used before video capsule endoscopy? A systematic review and meta-analysis *Eur J Gastroenterol Hepatol* 2014; 26: 137–145
- [11] Gkolfakis P, Tziatzios G, Dimitriadis GD et al. Meta-analysis of randomized controlled trials challenging the usefulness of purgative preparation before small-bowel video capsule endoscopy. *Endoscopy* 2018; 50: 671–683
- [12] Viazis N, Sgouros S, Papaxoinis K et al. Bowel preparation increases the diagnostic yield of capsule endoscopy: a prospective, randomized, controlled study. *Gastrointest Endosc* 2004; 60: 534–538
- [13] Wi JH, Moon JS, Choi MG et al. Bowel preparation for capsule endoscopy: a prospective randomized multicenter study. *Gut Liver* 2009; 3: 180–185
- [14] Rosa BJ, Barbosa M, Magalhaes J et al. Oral purgative and simethicone before small bowel capsule endoscopy. *World J Gastrointest Endosc* 2013; 67–73
- [15] Hookey L, Louw J, Wiepjes M et al. Lack of benefit of active preparation compared with a clear fluid-only diet in small-bowel visualization for video capsule endoscopy: results of a randomized, blinded, controlled trial. *Gastrointest Endosc* 2017; 85: 187–193
- [16] van Tuyl SA, den Ouden H, Stolk MF et al. Optimal preparation for video capsule endoscopy: a prospective, randomized, single-blind study. *Endoscopy* 2007; 39: 1037–1040
- [17] Franke A, Hummel F, Knebel P et al. Prospective evaluation of small bowel preparation with bisacodyl and sodium phosphate for capsule endoscopy. *World J Gastrointest Endosc* 2008; 14: 2061–2064
- [18] Postgate A, Tekkis P, Patterson N et al. Are bowel purgatives and prokinetics useful for small-bowel capsule endoscopy? A prospective randomized controlled study *Gastrointest Endosc* 2009; 69: 1120–1128
- [19] Park SC, Keum B, Hyun JJ et al. A novel cleansing score system for capsule endoscopy. *World J Gastrointest Endosc* 2010; 16: 875–880
- [20] Eliasziw M, Young SL, Woodbury MG et al. Statistical methodology for the concurrent assessment of interrater and intra-rater reliability: using goniometric measurements as an example. *Phys Ther* 1994; 74: 777–788
- [21] Davison ACHD. *Bootstrap methods and their application*. Cambridge: University Press; 1997
- [22] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159
- [23] Samaan MA, Shen B, Mosli MH et al. Reliability among central readers in the evaluation of endoscopic disease activity in pouchitis. *Gastrointest Endosc* 2018; 88: 360–369.e2
- [24] Hong-Bin C, Yue H, Su-Yu C et al. Evaluation of visualized area percentage assessment of cleansing score and computed assessment of cleansing score for capsule endoscopy. *Saudi J Gastroenterol* 2013; 19: 160–164
- [25] Chan RHL, Louw J et al. Reliability study to assess a measure of small bowel cleanliness in video capsule endoscopy. *Can J Gastroenterol Hepatol* 2008; 22: A341
- [26] Fireman Z, Paz D, Kopelman Y. Capsule endoscopy: improving transit time and image view. *World J Gastrointest Endosc* 2005; 11: 5863–5866
- [27] Dai N, Gubler C, Hengstler P. Improved capsule endoscopy after bowel preparation. *Gastrointest Endosc* 2005; 61: 28–31
- [28] Adler SN, Farkash S, Sompolinsky Y et al. A novel purgative protocol for capsule endoscopy of the small bowel produces better quality of visibility than 2 l of PEG: Timing is of the essence. *United European Gastroenterol J* 2017; 5: 485–490
- [29] Brotz C, Nandi N, Conn M et al. A validation study of 3 grading systems to evaluate small-bowel cleansing for wireless capsule endoscopy: a quantitative index, a qualitative evaluation, and an overall adequacy assessment. *Gastrointest Endosc* 2009; 69: 262–270
- [30] Esaki M, Matsumoto T, Kudo T et al. Bowel preparations for capsule endoscopy: a comparison between simethicone and magnesium citrate. *Gastrointest Endosc* 2009; 69: 94–101
- [31] Ito T, Ohata K, Ono A et al. Prospective controlled study on the effects of polyethylene glycol in capsule endoscopy. *World J Gastrointest Endosc* 2012; 18: 1789–1792
- [32] Oliva S, Cucchiara S, Spada C et al. Small bowel cleansing for capsule endoscopy in paediatric patients: a prospective randomized single-blind study. *Dig Liver Dis* 2014; 46: 51–55
- [33] Van Weyenberg SJ, De Leest HT, Mulder CJ. Description of a novel grading system to assess the quality of bowel preparation in video capsule endoscopy. *Endoscopy* 2011; 43: 406–411
- [34] Chen HB, Huang Y, Chen SY et al. Small bowel preparations for capsule endoscopy with mannitol and simethicone: a prospective, randomized, clinical trial. *J Clin Gastroenterol* 2011; 45: 337–341