Analysis of a grading system to assess the quality of small-bowel preparation for capsule endoscopy: in search of the Holy Grail

Authors Jatinder Goyal[1], Anshum Goel[1], Gerald McGwin[2,3], Frederick Weber[1,3]

Institutions
[1] Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA
[2] Department of Epidemiology, University of Alabama at Birmingham, Birmingham, Alabama, USA
[3] Department of Gastroenterology, University of Alabama at Birmingham, Birmingham, Alabama, USA

submitted 30. May 2014
accepted after revision 3. June 2014

Bibliography
DOI http://dx.doi.org/10.1055/s-0034-1377521
Published online: 16.7.2014
Endoscopy International Open 2014; 02: E183–E186

E-ISSN 2196-9736

Corresponding author
Jatinder Goyal, MD
Department of Medicine
University of Alabama
619 19th St. South Birmingham, AL 35205
USA
Fax: +1-205-934-9665
jatinaiims@gmail.com

Background: The diagnostic yield of capsule endoscopy is vulnerable to inadequate visualization related to residual bile or chyme remaining in the lumen despite intestinal lavage. It has been challenging to determine the optimal lavage preparation of the bowel and patient diet before capsule endoscopy, as well as the timing of the procedure, because no well-accepted, validated grading system for assessing the quality of intestinal lavage before capsule endoscopy is available. There remains no consensus on the reliability of qualitative, quantitative, or computer-derived assessments of the quality of preparation for capsule endoscopy. This study evaluates intra-observer and interobserver agreement for a previously validated scale.

Materials and methods: The digital images of 34 patients who underwent capsule endoscopy were independently reviewed by two blinded physicians according to a previously validated grading scale. One of the physicians reviewed and graded the patients a second time. The quality of the bowel luminal preparation was assessed with a qualitative parameter (fluid transparency) and a more quantitative parameter (mucosal invisibility) for each of three small-intestinal segments, and an overall small-bowel score for each parameter was assigned as well. A weighted kappa coefficient was used to calculate intra-observer (observer 1A and 1B) and interobserver (observer 1A and observer 2) agreement. A kappa value of 0.60 or more suggests strong agreement, 0.40 to 0.60 moderate agreement, and less than 0.40 poor agreement.

Results: The intra-observer weighted kappa index for both fluid transparency and mucosal visibility was 0.52, which is consistent with moderate agreement. The interobserver weighted kappa indices for fluid transparency and mucosal invisibility were 0.29 and 0.42, respectively, demonstrating suboptimal interobserver agreement. The individual segment interobserver kappa indices were better for mucosal visibility (0.52, 0.39, and 0.47 for small-bowel segments 1, 2, and 3, respectively) than for fluid transparency (0.18, 0.38, and 0.31).

Conclusions: The proposed grading scale for assessing the quality of preparation for capsule endoscopy has inadequate interobserver and intra-observer agreement. Capsule endoscopy preparation grading scales that focus more on quantitative than on qualitative assessment may demonstrate more reliable performance characteristics. Optimizing the quality of preparation and diagnostic yield of capsule endoscopy will first require the development of a well-validated grading scale.

Introduction

Capsule endoscopy is a noninvasive technology that is widely used to diagnose and monitor small-bowel pathology [1]. Current applications include obscure gastrointestinal bleeding, Crohn's disease, celiac disease, and polyposis syndromes [2]. Multiple factors (food material, air bubbles, bile, and blood) influence the quality of visualization and the diagnostic yield of capsule endoscopy. Individual variability in the small-bowel transit time can also influence outcomes in capsule endoscopy studies. Because capsule endoscopy does not inherently have suctioning and flushing capabilities, adequate bowel preparation plays a pivotal role in improving the quality of the images obtained and hence the diagnostic yield [3–5]. To date, consensus about the optimal strategy for bowel preparation before capsule endoscopy has not been achieved.

This lack of consensus is in part due to the lack of a standardized grading scale for small-bowel preparations in capsule endoscopy. Although multiple grading scales have been proposed [2,3,6–

10], wide variations have been noted among them, and few direct comparisons have been made [11]. Furthermore, most of these scales have not been validated at separate centers, so that their reliability across diverse clinical settings can be called into question. In this study, we sought to evaluate a previously validated grading scale [6] for bowel preparation before capsule endoscopy.

Materials and methods

A total of 34 capsule endoscopy images obtained from June 2011 to March 2012 in patients with definitive small-bowel pathology were selected from a tertiary care academic hospital. Both outpatients and inpatients were included for the study. Patients with a history of chronic narcotic use, prokinetics use (erythromycin, metoclopramide), gastroparesis, pseudo-obstruction, gastrectomy, or any other intestinal surgery were excluded from the study. The preparation before morning capsule endoscopy at our center consisted of a clear liquid diet following lunch the day before the examination; intestinal lavage with 2 L of polyethylene glycol was then followed by fasting after 10 p.m. the night before the study.

In this validation study, the 34 capsule endoscopy videos were prospectively read by two physicians (J. G. and F. W.) with extensive experience in reading them. One of the physicians (J. G.) reviewed and graded the readings a second time after an interval of 1 month (readings 1A and 1B). The videos were deidentified, and the reviewers were blinded to each patient's clinical records, demographics, and findings of the originally read capsule endoscopy video. They were also blinded to each other's capsule endoscopy interpretations. Both reviewers participated in a calibration session before capsule endoscopy analysis and interpretation.

Assessment of capsule endoscopy images

PillCam SB (Given Imaging, Yokneam, Israel) is a minimally invasive technology that captures images at the rate of two to six frames per second and transmits the images to an external data recorder. The images are then downloaded to a RAPID 5 workstation. The digital video images were analyzed for bowel preparation quality. The quality of both segmental and overall small-bowel preparation was graded individually by each reviewer.

Grading scale

We used the grading system previously validated by Esaki et al. to assess the quality of bowel luminal preparation [6]. The scale takes into account a qualitative measure (transparency of fluid) and a quantitative measure (degree of bowel mucosa visualized). The video images were equally divided into three segments after evaluation of the small-bowel transit time on the time counter. The fluid transparency of each segment was assessed and graded by reviewing the images at low speed (10 frames per second). Then, the degree of mucosal invisibility was assessed and graded by reviewing the images at maximum speed (40 frames per second) with concurrent manual inspection of individual frames. The grade of fluid transparency was determined according to the predominant grade in each segment. The grade of mucosal invisibility of each segment was determined by the proportion of the duration of the video image in which air bubbles or food residues disturbed more than 50% of visualization and interpretation. The grading system for the assessment of image quality is outlined in **Table 1**. Both reviewers rated each bowel segment individually, and a total small-bowel score was also calculated.

Table 1 Grading system for the assessment of image quality.

Grade	Description
Fluid transparency	
Grade I	Clear fluid that does not obscure vision
Grade II	Slightly dark fluid minimally obscuring vision
Grade III	Opaque fluid partly obscuring vision
Grade IV	Turbid fluid severely obscuring vision
Mucosal invisibility[1]	
Grade I	<5% of duration of video image with >50% bubbles or residues
Grade II	5%–15%
Grade III	15%–25%
Grade IV	>25%
Overall image quality[2]	
Grade A	3–5
Grade B	6–8
Grade C	9–12

[1] The percentage indicates the proportion of the duration of the video image in which air bubbles or food residues disturbed more than 50% of visualization and interpretation.

[2] The number indicates the sum of the grades in each small-intestinal segment.

Table 2 Weighted kappa indices for intra-observer and interobserver agreement.

	Weighted intra-observer kappa	Weighted inter-observer kappa
Fluid transparency		
Segment 1	0.55	0.18
Segment 2	0.47	0.38
Segment 3	0.39	0.31
Overall	0.52	0.29
Mucosal invisibility		
Segment 1	0.48	0.52
Segment 2	0.65	0.39
Segment 3	0.48	0.47
Overall	0.52	0.42

Statistical analysis

A weighted kappa coefficient was used as a measure of intra-observer (observer 1A and observer 1B) and interobserver (observer 1A and observer 2) agreement. A kappa value of 1 was considered to indicate perfect agreement, whereas a kappa value of 0 indicated agreement equivalent to chance. A kappa value of 0.6 or more suggested strong agreement, 0.4 to 0.6 moderate agreement, and 0.4 or less poor agreement [12]. SAS 9.1 was used for statistical analyses (SAS Institute, Cary, North Carolina, USA).

Results

A total of 34 capsule endoscopies were studied, with a total of 102 segments evaluated independently by the two clinicians. Kappa coefficients for fluid transparency and mucosal invisibility for individual segments and the overall small-bowel scores are summarized in **Table 2**.

Intra-observer reliability

Moderate intra-observer agreement was found between the two readings by the same reviewer (1A and 1B), with a weighted kappa value of 0.52 for the whole small bowel with respect to both fluid transparency and mucosal invisibility. The kappa coefficient

for fluid transparency showed moderate agreement for segments 1 and 2 (0.55 and 0.47, respectively) but poor agreement for segment 3 (0.39). There was moderate agreement between the two readings with respect to mucosal invisibility for segments 1 and 3 (0.48 each) and strong agreement for segment 2 (0.65).

Interobserver reliability

Interobserver reliability was worse than intra-observer reliability, with the overall kappa coefficient indicating poor agreement for fluid transparency (0.29) and moderate agreement for mucosal invisibility (0.42) for the entire small bowel. Weighted interobserver kappa coefficients for mucosal invisibility (0.52, 0.39, and 0.47) were better than those for fluid transparency (0.18, 0.38, and 0.31) for the individual segments.

Discussion

Capsule endoscopy is an extremely valuable diagnostic tool for detecting small-bowel mucosal disorders and has widely expanded the current scope of clinical practice. Although it is being increasingly adopted worldwide, it continues to be marred by certain limitations. Because it is incapable of suctioning, air insufflation, or luminal washing, images may contain air bubbles, intestinal secretions, bile, food residues, or blood. As a result, the diagnostic yield is lowered, and studies may be have to be repeated in as many as one-third of cases. The optimal small-bowel preparation remains a matter of debate; prokinetic agents like metoclopramide [13] and erythromycin [14], simethicone [7], and purgatives like polyethylene glycol [15] have been proposed in an attempt to improve bowel visualization. Various studies have suggested that improved small-bowel preparation could translate into an improved diagnostic yield of capsule endoscopy [16, 17]. However, to date no standardized procedure for small-bowel preparation for capsule endoscopy has been widely adopted. In addition, types of preparations, doses, and the timing of administration differ among various centers, so that direct comparisons are extremely difficult.

A meta-analysis by Rokkas et al. demonstrated that the diagnostic yield (odds ratio [OR], 1.813; $P=0.002$) and the quality of small-bowel visualization (OR, 2.113; $P=0.005$) were significantly better in patients whose bowel preparation consisted of purgatives than in those prepared with a clear liquid diet [18]. However, the study found no significant differences in capsule endoscopy completion rate, gastric transit time, or small-bowel transit time. A standardized grading scale for small-bowel preparation for capsule endoscopy is a prerequisite for developing techniques to improve the diagnostic yield.

Although a number of grading scales have been proposed to assess image quality [2, 3, 6 – 10], most of them are time-consuming, complicated, and difficult to apply in real-world clinical practice. They vary in the time and effort required to assess the bowel preparation, as well as in their definition of an adequate bowel preparation. Furthermore, most of the scales involve various degrees of scoring subjectivity and have not been well validated. **Table 3** outlines the scales and their parameters. Our study sought to evaluate one such novel and simple previously validated scoring system for capsule endoscopy bowel preparation [6]. This scale assessed bowel luminal preparation quality with qualitative (fluid transparency) and quantitative (mucosal invisibility) parameters and assigned a score for each parameter.

Table 3 Small-bowel preparation scales.

Study	Quantitative parameters	Correlation coefficient
Park et al. [2]	Proportion of visualized mucosa	0.80[1]
Viazis et al. [3]	Adequate vs. inadequate preparation	–
Esaki et al. [6]	Mucosal invisibility grade based on proportion of mucosa visualized	–
Albert et al. [7]	Mucosal visibility and intraluminal gas bubbles	0.78[2]
Niv et al. [8]	Proportion of visualized mucosa	–
Brotz et al. [9]	Quantitative index based on percentage of mucosa visualized, fluid and debris, bubbles, bile/chyme staining, and brightness	0.47 – 0.52[2]
Van Weyenburg et al. [10]	Computed assessment of cleansing (CAC) score based on color intensities	0.68[3]

[1] Intraclass coefficient (ICC).
[2] Kappa coefficient.
[3] Spearman's rho.

Despite a mutual calibration session, our results demonstrate that this grading scale offers inadequate performance characteristics and differ from the results of Esaki et al., who found more favorable interobserver variability. Their study found strong interobserver agreement for fluid transparency in all three small-intestinal segments ($r=0.88$, 0.77, and 0.81, respectively). In our study, the interobserver kappa coefficient for the qualitative parameter of fluid transparency was 0.29, which is consistent with poor overall agreement between the two investigators. However, we did find that the kappa value for the quantitative parameter of mucosal invisibility was better, at 0.42. Although the concordance is not as strong as that reported from the original study, it still has significant implications. Our results suggest that grading scales that employ more quantitative parameters may have lower rates of interobserver variability. This is consistent with the results of Brotz et al., who showed that a quantitative index was better than a qualitative index for interobserver and intra-observer reliability [9]. Quantitative parameters employ easily measurable characteristics that can be reliably assessed by investigators. These parameters minimize the level of subjectivity in the scales and thus improve their performance characteristics. Scales relying on qualitative interpretation may suffer from poor concordance.

Prior studies have used various parameters along with complicated scoring systems. The dichotomous (adequate vs. inadequate) grading system proposed by Viazis et al. [3] showed impressive concordance between investigators but may be inadequate for describing the differences in viewing quality associated with various diets, lavage regimens, and timing of capsule endoscopy preparations. An alternative scoring system employed by Park et al. involves viewing representative frames at 5-minute intervals and scoring them on two parameters: proportion of visualized mucosa and degree of obscuration [2]. The authors reported excellent interobserver and intrapatient concordance. This scale has distinct advantages of being simple, efficient, and easy to use on an everyday basis. Brotz et al. validated three parameters for small-bowel cleansing: a quantitative index (0 – 10); a qualitative evaluation (poor, fair, good, excellent); and an overall adequacy assessment (inadequate, adequate) [9]. Rates of both intra-observer and interobserver reliability were higher for the quantitative in-

dex and overall adequacy assessment than for the qualitative evaluation, and a dichotomized quality evaluation of excellent/good vs. fair/poor had moderate to substantial intra-observer and interobserver reliability. Similarly, our study provides further evidence that quantitative parameters for scales might provide a better assessment of small-bowel preparation for capsule endoscopy. In another study, Van Weyenberg et al. proposed a novel computed assessment of small-bowel mucosal visibility based on the ratio of the color intensities of the red and green channels of the tissue color bar [10]. Whether automated scoring will prove superior requires further investigation.

Our study has multiple strengths. We employed a previously validated scale that was simple to use. The reviewers underwent a calibration session before rating the studies. They were blinded to each other's ratings to eliminate bias.

A limitation of our study is that it assessed only one of the previously validated scales published in the literature. Our findings could not confirm the findings of Esaki et al. and underscore the need for validation by multiple investigators in various clinical settings before widespread acceptance can be achieved.

Conclusion

The proposed scoring system for grading capsule endoscopy image quality demonstrates suboptimal interobserver and intra-observer performance characteristics. However, our results for interobserver agreement on the quantitative parameter of mucosal invisibility are encouraging. Capsule endoscopy preparation scoring systems that are based on quantitative assessment will likely be more reliable than those based on qualitative parameters. Further studies are needed to establish a reliable grading scale for small-bowel preparation. This will be a prerequisite to defining the optimal diet, lavage regimen, and timing of lavage before capsule endoscopy, so that a maximal diagnostic yield can be achieved.

Competing interests: None

References

1 *Ladas SD*, *Triantafyllou K*, *Spada C* et al. European Society of Gastrointestinal Endoscopy (ESGE): recommendations (2009) on clinical use of video capsule endoscopy to investigate small-bowel, esophageal and colonic diseases. Endoscopy 2010; 42: 220 – 227

2 *Park SC*, *Keum B*, *Hyun JJ* et al. A novel cleansing score system for capsule endoscopy. World J Gastroenterol 2010; 16: 875 – 880

3 *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

4 *Koornstra JJ*. Bowel preparation before small bowel capsule endoscopy: what is the optimal approach? Eur J Gastroenterol Hepatol 2009; 21: 1107 – 1109

5 *Park SC*, *Keum B*, *Seo YS* et al. Effect of bowel preparation with polyethylene glycol on quality of capsule endoscopy. Dig Dis Sci 2011; 56: 1769 – 1775

6 *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

7 *Albert J*, *Gobel CM*, *Lesske J* et al. Simethicone for small bowel preparation for capsule endoscopy: a systematic, single-blinded, controlled study. Gastrointest Endosc 2004; 59: 487 – 491

8 *Niv Y*, *Niv G*, *Wiser K* et al. Capsule endoscopy – comparison of two strategies of bowel preparation. Aliment Pharmacol Ther 2005; 22: 957 – 962

9 *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

10 *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

11 *Villa F*, *Signorelli C*, *Rondonotti E* et al. Preparations and prokinetics. Gastrointest Endosc Clin N Am 2006; 16: 211 – 220

12 *Viera AJ*, *Garrett JM*. Understanding interobserver agreement: the kappa statistic. Family Med 2005; 37: 360 – 363

13 *Selby W*. Complete small-bowel transit in patients undergoing capsule endoscopy: determining factors and improvement with metoclopramide. Gastrointest Endosc 2005; 61: 80 – 85

14 *Leung WK*, *Chan FK*, *Fung SS* et al. Effect of oral erythromycin on gastric and small bowel transit time of capsule endoscopy. World J Gastroenterol 2005; 11: 4865 – 4868

15 *Kotwal VS*, *Attar BM*, *Gupta S* et al. Should bowel preparation, antifoaming agents, or prokinetics be used before video capsule endoscopy? A systematic review and meta-analysis Eur J Gastroenterol Hepatol 2014; 26: 137 – 145

16 *Ito T*, *Ohata K*, *Ono A* et al. Prospective controlled study on the effects of polyethylene glycol in capsule endoscopy. World J Gastroenterol 2012; 18: 1789 – 1792

17 *Dai N*, *Gubler C*, *Hengstler P* et al. Improved capsule endoscopy after bowel preparation. Gastrointest Endosc 2005; 61: 28 – 31

18 *Rokkas T*, *Papaxoinis K*, *Triantafyllou K* et al. Does purgative preparation influence the diagnostic yield of small bowel video capsule endoscopy? A meta-analysis Am J Gastroenterol 2009; 104: 219 – 227