Validation of the computed assessment of cleansing score with the Mirocam® system

Ana Ponte, Rolando Pinho, Adélia Rodrigues, Joana Silva, Jaime Rodrigues and João Carvalho

Department of Gastroenterology. Centro Hospitalar Vila Nova de Gaia/Espinho. Gaia, Portugal

ABSTRACT

Background and aims: A computed assessment of cleansing (CAC) score was developed to objectively evaluate small-bowel cleansing in the PillCam capsule endoscopy (CE) system and to overcome the subjectivity and complexity of previous scoring systems. Our study aimed to adapt the CAC score to the Mirocam® system, evaluate its reliability with the Mirocam® CE system and compare it with three validated subjective grading scales.

Patients and methods: Thirty CE were prospectively and independently reviewed by two authors who classified the degree of small-bowel cleanliness according to a quantitative index, a qualitative evaluation and an overall adequacy assessment. The authors were blinded for the CAC score of each CE, which was calculated as (mean intensity of the red channel)/[mean intensity of the green channel] - 1) x 10. The mean intensities of the red and green channels of the small-bowel segment of the “Map View” bar in the Miroview Client® were determined using the histogram option of two photo-editing software.

Results: There was a strong agreement between both CE readers for each of the three subjective scales used. The reproducibility of the CAC score was excellent and identical results were obtained with the two photo-editing software. Regarding the comparison between the CAC score and the subjective scales, there was a moderate-to-good agreement with the quantitative index, qualitative evaluation and overall adequacy assessment.

Conclusions: CAC score represents an objective and feasible score in the assessment of small-bowel cleansing in the Mirocam® CE system, and could be used per se or as part of a more comprehensive score.

Key words: Capsule endoscopy. Small-bowel cleansing scales.

INTRODUCTION

Capsule endoscopy (CE) was developed in 2001 and represents a safe and non-invasive diagnostic procedure which allows complete small-bowel mucosal assessment and guides further diagnostic and therapeutic approaches (1-6). Currently, CE has an essential role in the management of patients with suspected small-bowel diseases, including obscure gastrointestinal bleeding (OGIB), Crohn’s disease, small-bowel tumors, polyposis syndromes and celiac disease (1,7-11). Although previous studies report a higher diagnostic yield compared with other methods including push enteroscopy, radiologic procedures or angiography, CE has some limitations (5-7,12-16). Two factors which negatively impair the diagnostic yield of CE are the presence of intestinal debris and air bubbles and the failure to complete an entire small-bowel examination (2,11,17,18). The evaluation of the quality of small-bowel cleansing is necessary to assess the reliability of the findings in CE (8,17). Moreover, a consensus regarding the need for intestinal preparation for CE remains to be established (9,15,18). The presence of multiple grading scales for small-bowel preparation in CE which are time-consuming and complicated contributes to the difficulty in comparing different small-bowel cleansing regimens and to apply them to clinical practice (9,17). Furthermore, most of them are not validated and are based on subjective parameters (9-11,14,17-21). Van Weyenberg et al. developed a computed assessment of the cleansing (CAC) score, based on objective measurements of color intensities in red and green channels of the tissue color bar in the Rapid Reader® of the PillCam® CE system (PillCam, Medtronic plc. Dublin, Ireland), thus avoiding subjectivity (17). The authors hypothesized that if the tissue color bar, which comprises the summary of all CE images, was converted to the red-green-blue mode (RGB), the ratio between the mean intensity of the red and green channels could be used as a measure of enteric cleanliness. Therefore, areas of adequate mucosal visibility should be associated with high values of red intensity and low values of green intensity. Conversely, areas with high amounts of intestinal debris should be associated with low values of red intensity and high values of green intensity (17).

The aims of this study were to adapt the CAC score (17) to the Mirocam® system (Mirocam, IntroMedic, Seoul,
Korea), to evaluate its reliability with the Mirocam® CE system and to compare this automated scale with three previously validated subjective grading scales (8).

PATIENTS AND METHODS

Capsule endoscopies

Capsule endoscopies were included in the study if the patients were above 18 years old and the capsule reached the cecum within the recording time. Patients with diabetes mellitus, thyroid pathology or taking drugs known to interfere with gastrointestinal motility were excluded from the study. Furthermore, patients with active bleeding confirmed during CE were also excluded from this study. All patients gave informed consent for the CE examination. Thirty capsule endoscopies fulfilling the inclusion and exclusion criteria, performed between November 2014 and June 2015, were selected consecutively.

All CE were performed at the same center using the Mirocam® CE system (Mirocam, IntroMedic, Seoul, Korea). This CE records images at a rate of 3 frames per second (fps) that are transmitted through human body communication technology to an external recorder. The CE protocol consists of capsule ingestion with a glass of water at 8 a.m. after an overnight fast without prior bowel preparation, real-time views at 1 h and 2 h post-ingestion, administration of 10 mg metoclopramide if the capsule remains in the stomach after 2 h, light diet 4h after capsule ingestion, and normal daily activities on an outpatient basis during the day; the recorder is removed 12 h after capsule ingestion (or earlier, if real-time viewing confirms that the capsule has already reached the colon).

All CE were reviewed by two authors (AP, RP) who confirmed the first duodenal image and the first cecal image. All remaining findings were deleted from the CE videos. The resulting video files were saved and numbered with sequential numbers ranging from 1 to 30. A dedicated database including the number of the CE, age and sex of the patient, the indication for CE and fields for the subjective scales and the computer scores was created.

Subjective scales

All CE videos were reviewed independently by two authors (AP, RP) at a fixed rate of 40 fps, in Single View (Miroview Client®), with the purpose of grading the level of cleanliness according to 3 subjective grading scales previously described (8): a) quantitative index (QI), which grades the level of cleanliness with a score ranging from 0 to 10; b) qualitative evaluation (QE), which grades the degree of cleanliness as excellent, good, fair and poor; and c) overall adequacy assessment (OAA), which is simply graded as adequate or inadequate. Both readers were blinded to the results of the CAC score.

Computed assessment of cleansing score

Besides Single View, Dual View and Quad View, the Mirocam® reading software (Miroview Client®) has another function named Map View which displays a bar containing a representation of all images recorded by the CE. Although this bar can be zoomed, without zoom the bar is similar to the tissue color bar in the Rapid Reader® in the PillCam® CE system (PillCam, Medtronic plc. Dublin, Ireland). The small-bowel in the Map View bar is automatically circumscribed by a blue line according to the images previously marked as the first duodenal and cecal images. An image of the Map View bar was created using free open source software (ShareX®, available at http://getsharex.com) (Fig. 1A). The resulting image was imported into 2 photo editing software (proprietary software, Photoshop® [Photoshop CS2, version 9.0.2, Adobe Systems Inc., San Jose, California, USA], and free open source software, GIMP®, GNU Image Manipulation Program, available at https://www.gimp.org/]). The images were cropped using both softwares to include only the segment corresponding to the small-bowel. The histogram function of Photoshop® and GIMP® was then used to measure the mean intensity value of the green and red channels (Fig. 1B). These values were entered into the database and the computer scores were determined according to the formula published by Weyenberg et al. (17): ([mean intensity of the red channel]/[mean intensity of the green channel] - 1) x 10 for both Photoshop® (P-CAC) and GIMP® (G-CAC).

Statistical analysis

Descriptive results are presented as percentages, means and standard deviation of the mean. The correlation between the computer scores obtained with Photoshop® and GIMP® and between the computer scores and the 10 points QI was performed using Spearman’s correlation coefficient. The computer scores and the 4 grades QE were compared using the intra-class correlation coefficient (ICC). The computer scores were compared according to the OAA using Student’s t test.

The level of agreement between both readers was evaluated using the ICC for the QI and the QE and the unweighted kappa test for the OAA.

Fig 1. A. Image depicting the print screen of the entire tissue color bar of the Map View. B. Histogram function of GIMP® used to measure the mean intensity value of the green and red channels of the small bowel segment of the tissue color bar.
Results were considered to be significant when p < 0.05. The Statistical Package for Social Sciences version 20.0 (IBM Corp., Armonk, New York, USA) was used for data entry and data analysis.

RESULTS

The 30 CE videos referred to 30 patients, 66.7% males with a mean age of 49.3 (±17) years old. The main indications for CE were OGIB and suspected Crohn’s disease (Table I). The mean small-bowel transit time was 272.2 (±71.3) minutes.

CAC scores

The mean value of the CAC score was 4.93 (±1.36) using Photoshop® and 4.95 (±1.36) using GIMP®. An excellent correlation between the 2 scores was found with a Spearman’s Rho of 1.00, p < 0.001.

Quantitative index

The mean grade for the QI was 7.47 (±1.74) for reader 1 and 7.9 (±1.67) for reader 2 (Table II). An excellent inter-reader reliability with an ICC of 0.92 (CI 95% = 0.84-0.96) was found.

A good correlation between the QI and the computed scores was found with a Spearman’s Rho of 0.5, p = 0.005 for reader 1; 0.48, p = 0.008 for reader 2; and 0.49, p = 0.006 for the mean of both readers using P-CAC, and 0.49, p = 0.006 for reader 1; 0.47, p = 0.009 for reader 2 and 0.49, p = 0.007 for the mean of both readers using G-CAC (Fig. 2).

Qualitative evaluation

The scores obtained with the QE scale are detailed in table III. An excellent concordance between both readers with an ICC of 0.83 (CI 95% = 0.68-0.91) was found.

A fair correlation between QE and the computed scores was found with an ICC of 0.43 (CI 95% = 0.089-0.68, p = 0.008) for reader 1 and 0.45 (CI 95% = 0.11-0.70, p = 0.006) for reader 2 using P-CAC and 0.43 (CI 95% = 0.088-0.68, p = 0.008) and 0.45 (CI 95% = 0.11-0.69, p = 0.008), respectively, using G-CAC.

Overall adequacy assessment

The degree of cleanliness was considered as adequate in 86.7% of cases by reader 1 and 90% of cases by reader 2. An excellent inter-reader reliability was found, with an unweighted kappa of 0.84. The mean values of P-CAC for adequate/inadequate levels of cleanliness were 5.16/3.45 (p = 0.02) for reader 1 and 5.07/3.66 for reader 2 (p = 0.09); the mean values for G-CAC were 5.18/3.47 (p = 0.02) and 5.09/3.68 (p = 0.09), respectively (Fig. 3).

Table I. Demographic data of patients and indications for capsule endoscopy performance

<table>
<thead>
<tr>
<th>Demographic data</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male/Female</td>
<td>20 (66.7)/10 (33.3)</td>
</tr>
<tr>
<td>Mean age +/- S.D. (years)</td>
<td>49.3 +/- 17</td>
</tr>
<tr>
<td>Indication for CE</td>
<td></td>
</tr>
<tr>
<td>OOGIB - Anemia</td>
<td>14 (46.7)</td>
</tr>
<tr>
<td>Suspected Crohn’s disease</td>
<td>13 (43.3)</td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>2 (6.7)</td>
</tr>
<tr>
<td>OOGIB - positive FOBT</td>
<td>1 (3.3)</td>
</tr>
</tbody>
</table>

CE: Capsule endoscopy; OOGIB: Occult obscure gastrointestinal bleeding; FOBT: fecal occult blood test.

Table II. Results of the 10-point quantitative scale for the evaluation of the small bowel cleansing according to reader 1 and reader 2

<table>
<thead>
<tr>
<th>Reader 1 (No. [%])</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 (3.3)</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 (6.7)</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td></td>
<td>1</td>
<td></td>
<td>0</td>
<td>0</td>
<td>3 (10)</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5 (16.7)</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>7 (23.3)</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6 (20)</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>6 (20)</td>
</tr>
<tr>
<td>Total</td>
<td>1 (3.3)</td>
<td>1 (3.3)</td>
<td>2 (6.7)</td>
<td>3 (10)</td>
<td>6 (20)</td>
<td>10 (33.3)</td>
<td>3 (10)</td>
<td>4 (13.3)</td>
<td>30</td>
</tr>
</tbody>
</table>
DISCUSSION

CE is a safe, non-invasive and extremely valuable tool in the diagnosis of small-bowel pathologies (9). Nevertheless, the quality of mucosal visualization and the diagnostic yield of CE may be impaired due to incomplete examination and poor intestinal preparation with air bubbles, bile, intraluminal fluid and debris (2,7,9,15,19,21,22). Moreover, CE is time consuming and expensive and has inherent limitations including uncontrolled movements, unidirectional field of view, incapability of air inflation, fluid suction or washing of the small-bowel mucosa (3,8,10,14,20-23). Therefore, it is essential to optimize intestinal preparation to improve mucosal assessment in order to prevent missing lesions and the need to repeat CE, which may occur in up to 33% of the cases (8,10,20,22).

Currently, the bowel preparation method recommended for CE consists of a clear liquids diet and a 12-hour fast (2,7,8,10,11,13,14,19,24,25). This recommendation probably resulted from the experience with push enteroscopy, which evaluates the proximal small-bowel usually 80-120 cm beyond the ligament of Treitz (7,26). Nevertheless, using this approach, the quality of CE images is affected in up to 30% of cases with residual intestinal contents, mainly in the distal part of the small-bowel, thus limiting mucosal visualization and possibly the diagnostic yield (5,11,13,14,21,25). As an adequate preparation improves adequacy of mucosal assessment, many stud-
ies have focused on the ideal approach to clean the small intestine before CE (5,7,10-15,18,19,21,25). Several bowel preparation methods have been described, namely the use of purgatives including polyethylene glycol, sodium phosphate, magnesium citrate or mannitol; simethicone; or prokinetics (8,9,15,17,27). Due to the heterogeneity of protocols regarding the type of preparation, dosage and time of administration, results are conflicting and a consensus regarding the efficacy, tolerability and best approach for intestinal cleansing before CE is lacking (7,9,10,12,14,15,18,19,21). Nevertheless, small-bowel preparation before CE seems to improve the quality of visualization of the mucosa and the diagnostic yield, without any difference regarding CE completion rate, gastric transit time and small-bowel transit time (2,27).

The absence of a standardized grading scale for small-bowel preparation in CE also contributes to the lack of consensus regarding the optimal small bowel cleansing method (9). Currently, a validated scale that is universally accepted for grading small-bowel cleansing is lacking. Moreover, most of them are based on subjective parameters (9,10,11,14,17-21). Weyenberg developed and validated a computed quantitative score for the PillCam® CE system, based on objective measurements of color intensities in red and green channels of the tissue color bar of CE, therefore avoiding subjectivity (17). Furthermore, the authors suggested that the incorporation of this score into the CE reading software would result in a fully automated score, as the manual selection of the tissue color bar segments and analysis in a photo-editing software would be redundant (17). This aspect would overcome all the disadvantages of the previous scoring systems, namely subjectivity, complexity and lengthiness. Taking into account all the advantages

![Box plot graphics showing the relation between the computed assessment of cleansing score (y-axis) using Photoshop® (P-CAC) and GIMP® (G-CAC) and the overall adequacy assessment (OAA) for readers 1 and 2 (x-axis).](image-url)
of CAC score, our study aimed to evaluate this novel and objective scale for CE bowel preparation using another CE platform: Mirocam®.

In our study, there was a strong agreement between both CE readers for each of the three subjective scales used. Namely, the inter-reader reliability for the QI was 0.92 (95% CI 0.84-0.96), 0.83 (95% CI 0.68-0.91) for the QE and 0.84 for the OAA. These findings were similar to those of Brotz et al., in which the QI had the greatest reliability, the reliability for the OAA was in the moderate range, whilst the QE performed less well (8). Quantitative scales provide parameters that are assessed more uniformly thus reducing subjective interpretation and providing a better evaluation of small-bowel preparation (8,9).

The reproducibility of the CAC score was excellent as the results of the two different photo-editing software were identical, resulting in an intra-test reliability of 1.0 (p < 0.001). Regarding the comparison between the CAC score and the subjective scales, there was a moderate-to-good agreement with the QI, QE and OAA. Our results were slightly inferior to those of Van Weyenberg (17) but still significant, and reinforce the feasibility of the CAC score in the assessment of the intestinal preparation in CE systems other than the PillCam®. The use of a visual analog scale which cannot differentiate between bubbles, debris and bile may explain the suboptimal results achieved. Nevertheless, the results are encouraging and may contribute to generate more research in order to improve this score and even achieve an optimal computed score. A fully automated cleansing score may overcome the subjectivity, complexity and lengthiness inherent to currently available scales. Similar to colonoscopy, a CE report should include an evaluation of small-bowel cleansing, in order to allow the assessment of the reliability of CE findings and the comparison of different bowel cleansing methods (8,17).

As proposed by Van Weyenberg et al., the CAC score should be validated in patients with different types of bowel preparation (17). In contrast to their study, in which patients received 2l of polyethylene glycol solution, in our study patients underwent CE after an overnight fast.

The main potential drawbacks of this study include the absence of a specific assessment of the CAC score in the evaluation of enteric preparation in the distal half of the small-bowel and the lack of analysis of the correlation of the quality of intestinal cleansing and the diagnostic yield of CE. Moreover, the three subjective scales were concomitantly graduated after each CE reading, which could influence the results, namely the concordance among the three scores. Nevertheless, this aspect was not evaluated in this study.

In conclusion, the CAC score also represents an objective and feasible score in the assessment of small-bowel cleansing in CE using the Mirocam® system, and could be used per se or as part of a more comprehensive score.

REFERENCES


