Is virtual chromoendoscopy useful in the evaluation of subtle ulcerative small-bowel lesions detected by video capsule endoscopy?

Authors

Mihai Rimbaș1,4, Lucian Negreanu2,3, Lidia Ciobanu4, Andreea Benguș1, Cristiano Spada3, Cristian Răsvan Băicuș2,4, Guido Costamagna3

Institutions

Institutions are listed at the end of article.

Background: The identification of subtle small-bowel mucosal lesions by video capsule endoscopy (VCE) can be challenging. Virtual chromoendoscopy techniques, based on narrowing the bandwidth of conventional white light endoscopic imaging (WLI), were developed to improve the analysis of mucosal patterns. However, data on the already-implemented Flexible spectral Imaging (or Fujinon Intelligent) Color Enhancement (FICE) software application in VCE are limited.

Materials and methods: An evaluation of 250 difficult-to-interpret small-bowel ulcerative and 50 artifact lesions selected from 64 VCE recordings was conducted by four experienced VCE readers in two steps: initially as WLI, then with the addition of all available virtual chromoendoscopy pre-sets (FICE 1, 2, and 3 and Blue mode). The readers labeled them as real or false ulcerative lesions and rated the usefulness of each of the pre-sets.

Results: Between the first (WLI-only) and second (virtual chromoendoscopy-aided) readings, in terms of accuracy there was a global 16.5 % (95 % confidence interval [95 %CI] 13.6 – 19.4 %) improvement (P < 0.001), derived from a 22 % (95 % CI 18.9 – 25.1 %) improvement in the evaluation of true ulcerative images (P < 0.001) and an 11 % (95 % CI 4.1 – 17.7 %) decrease in the evaluation of false ulcerative ones (P = 0.003). The FICE 1 and 2 pre-sets were rated as most useful.

Conclusion: The application of virtual chromoendoscopy for VCE is useful to better categorize difficult-to-interpret small-bowel mucosal ulcerative lesions. However, care must be taken, and individual images should be evaluated only as part of a sequence in a recording because the technology can also mistakenly guide to the incorrect interpretation of artifacts as ulcerative lesions.

Introduction

Video capsule endoscopy (VCE) examination of the small bowel is a noninvasive investigation that is able to capture images with excellent resolution, providing a diagnostic macroscopic evaluation of this organ. Technically easy and well tolerated, and carrying a low risk for adverse events, VCE examination is currently regarded as the first-line diagnostic tool in the context of small-bowel disease [1]. However, the examination is not perfect. Despite the fact that it has been constantly improved in terms of systems and protocols, such as angle of view, analysis software, and the use of prokinetic agents or bowel cleansing. Among its drawbacks, the detection of small differences in mucosal hue or pattern, which is necessary for the identification of subtle small-bowel lesions, can sometimes be challenging. To overcome this problem and enhance the contrast between adjacent mucosal areas, virtual chromoendoscopy techniques were developed. Of the different virtual chromoendoscopy technologies available, Flexible spectral Imaging (or Fujinon Intelligent) Color Enhancement (FICE) and Blue mode filtering modules have been incorporated into the reading software (RAPID) of the PillCam video capsule (Given Imaging Ltd., Yoqneam, Israel). The benefits expected from this technology, in terms of improving the identification or characterization of mucosal lesions, aside from being presented in case reports [2, 3], have thus far been assessed in only a few small comparative studies [4 – 7]. These have yielded scanty and discordant data that nonetheless suggest a potential role for FICE or Blue mode filtering in small-bowel VCE. However, the true added benefit of this technology can be underestimated if only clearly visible mucosal lesions are evaluated. Therefore, we undertook to test the technology by assessing less clearly visualized pathological changes of the mucosa.
The study consisted of a retrospective evaluation of 64 VCE recordings conducted in a tertiary care referral center and teaching hospital as part of our prospective database of patients collected from October 2007 to December 2013. In most cases, mild inflammatory involvement of the small-bowel mucosa was expected because the cohort included patients with seronegative spondyloarthritides who were taking nonsteroidal anti-inflammatory drugs (NSAIDs) [8]. The original study protocol was approved by the institutional ethics committee.

**Video capsule endoscopy procedure**

All the patients swallowed a PillCam video capsule (SB 2 or SB 2L; Given Imaging) after fasting for a minimum of 12 hours and ingesting a routine polyethylene glycol (PEG)-based bowel preparation (Endofalk; Dr. Falk Pharma GmbH, Freiburg, Germany), with 2L taken the evening before and 1L the morning of the procedure; the last sip was taken at least 1 hour before VCE ingestion. An 80-mg dose of simethicone (Espumisan L; Berlin-Chemie AG, Berlin, Germany) was given orally with a small quantity of water 15 to 20 minutes before initiation of the VCE examination. An 80-mg dose of simethicone (Espumisan L; Berlin-Chemie AG, Berlin, Germany) was given orally with a small quantity of water 15 to 20 minutes before initiation of the VCE examination. Clear fluids and a light meal were allowed at least 2 and 4 hours after video capsule ingestion, respectively.

**Selection of study images**

All the available VCE studies were carefully reviewed with RAPID 8 software and white light endoscopic imaging (WLI) to identify all the mucosal defects that might represent ulcerative lesions, defined as lesions with a pale or yellow base and a red or pink collar [9]. A single experienced reader (M.R.) conducted the review in a room with dimmed lights while using RAPID 8 software with the Image Adjustment function “off” and the following predefined settings: sharpness 1, brightness 0, and color 0. The process involved the identification of all lesions or artifacts that might be appropriate for the present study. In this regard, an ulcerative image had to appear on three different images in a recording and to present a characteristic red halo in order to be considered a true ulcerative image, and an artifact had to be clearly seen as an artifact (with no characteristic red halo, representing mainly white or yellow dirty luminal content) on at least three successive images. When it was uncertain whether a sequence of images represented an ulcerative lesion or an artifact, that sequence was abandoned, and the reader proceeded to evaluate the next sequence of images.

Two different sets of images were thus identified and selected. The first set consisted of what were most likely ulcerative lesions, carefully selected from a succession of images as the least representative visualization (but still suggestive of an ulcerative lesion) of an unequivocally confirmed erosion or ulceration; these comprised small or shallow mucosal defects, erosions lacking a clear rim of erythema or located marginally or distantly in the field of view, lesions with poor image quality due to opaque luminal content, image motion artifacts, or a combination of them (Fig. 1). The other set consisted of artifact images that mimicked ulcerative lesions but in fact represented dirty luminal content or particular arrangements of the mucosal villi. Moreover, for the selection of true ulcerative images, if it was not possible to obtain a less clear visualization (i.e., make it “subtle”) and for the evaluation of false ulcerative lesions, if it was not possible to find an image that could mimic a true ulcerative image, those lesions were also abandoned. The purpose was to form a database of difficult-to-interpret suspected ulcerative images and to be almost sure about what these images represented. The FICE (pre-sets 1, 2, and 3) and Blue mode correspondents of each and every one of the selected images were then created by using the Rapid software adjuncts (Fig. 2). All of these de-identified images were then mixed in random order to form a single set of true and false ulcerative still WLI pictures alternating with their virtual chromoendoscopic correspondents. The gold standard in the present study, although subjective, was the designation at initial selection (i.e., true ulcerative or not).

**Evaluation of the study images**

In a second phase, a multicenter evaluation of the selected images was done. The evaluations were done separately by four experienced VCE readers (C.S., L.N., L.C., and A.B.), blinded to the original findings, in two steps: initially as WLI, then with the addition of all available virtual chromoendoscopy settings (FICE pre-sets 1, 2, and 3 and Blue mode). The investigators asserted the presence or absence of the ulcerative lesions and rated the usefulness of each of the virtual chromoendoscopic pre-sets. All of the VCE evaluators in the present study are experienced endoscopists. C.S. is an internationally recognized expert in the field of VCE. M.R., L.C., and L.N. have individually reviewed more than 200 VCE studies, and A.B. has reviewed more than 100 VCE studies. The primary outcome measure of the study was a comparison of the accuracies with which the lesions were categorized in the first and second evaluations (McNemar’s test). Interobserver agreement for each reading was also calculated (Fleiss’s kappa test). Statistical analyses were carried out with SPSS version 16.0 for Windows (IBM, Armonk, New York, USA) and WinPEPI 11.39.
Results

Image selection

Initially, 310 true ulcerative images were identified. Of those, only 250 were finally retained because they were considered the most representative (i.e., difficult-to-interpret lesions) for the purpose of the present study. As for the set of false ulcerative images, from the 62 initially selected, only the 50 most suggestive of an ulcerative lesion were finally kept. Thus, a set of 300 WLI still pictures was created (250 true ulcerative and 50 false ulcerative lesions), along with their virtual chromoendoscopic correspondents.

Overall image evaluation

One of the four evaluators could not classify as true or false two of the images in the WLI set (both true ulcerative). All of the remaining images were rated as true or false by all the investigators in the two performed readings. At the first (WLI only) evaluation, the four observers rated 543 of 1198 images as true ulcerative and 655 as false ulcerative. In the second step, evaluating the same images while aided by virtual chromoendoscopy (FICE pre-sets 1, 2, and 3 and Blue mode), they designated 787 of 1200 images as true ulcerative and 413 as false ulcerative. Between the first and second readings, in terms of accuracy, there was a statistically significant improvement in overall (WLI and chromoendoscopy) image evaluation. The performance characteristics of the four evaluators are shown in Table 1. As the results indicate, the use of virtual chromoendoscopy improved the accuracy of differentiation in both true and false ulcerative lesions.

Table 1 Performance characteristics in the evaluation of the study images for the first (white light endoscopic imaging only) and second (chromoendoscopy-aided) readings.

| a) Overall image evaluation | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Observer | Accuracy with WLI only, % | Accuracy with chromoendoscopy, % | Improvement in accuracy, % [95% CI] | P value |
| 1 | 48.7 | 62.0 | 13.3 [7.5, 19.0] | < 0.001* |
| 2 | 59.0 | 76.0 | 17 [10.1, 23.6] | < 0.001* |
| 3 | 48.0 | 76.2 | 28.2 [21.7, 34.3] | < 0.001* |
| 4 | 59.0 | 66.7 | 7.7 [3.6, 11.7] | < 0.001* |
| Globally | 53.7 | 70.2 | 16.5 [13.6, 19.4] | < 0.001* |

b) True ulcerative image evaluation

<table>
<thead>
<tr>
<th>Observer</th>
<th>Accuracy with WLI only, %</th>
<th>Accuracy with chromoendoscopy, %</th>
<th>Improvement in accuracy, % [95% CI]</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40.8</td>
<td>59.2</td>
<td>18.4 [12.0, 24.5]</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>2</td>
<td>58.4</td>
<td>82.4</td>
<td>24 [17.0, 30.7]</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>3</td>
<td>42.3</td>
<td>77.8</td>
<td>35.5 [28.4, 41.9]</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>4</td>
<td>56.0</td>
<td>66.4</td>
<td>10.4 [5.9, 14.8]</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Globally</td>
<td>49.4</td>
<td>71.4</td>
<td>22 [18.9, 25.1]</td>
<td>&lt; 0.001*</td>
</tr>
</tbody>
</table>

c) False ulcerative image evaluation

<table>
<thead>
<tr>
<th>Observer</th>
<th>Accuracy with WLI only, %</th>
<th>Accuracy with chromoendoscopy, %</th>
<th>Decrease in accuracy (%) [95% CI]</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>88</td>
<td>76</td>
<td>12 [-0.6, 23.6]</td>
<td>0.109</td>
</tr>
<tr>
<td>2</td>
<td>62</td>
<td>44</td>
<td>18 [-1.3, 35.9]</td>
<td>0.106</td>
</tr>
<tr>
<td>3</td>
<td>76</td>
<td>68</td>
<td>8 [-4.6, 20.0]</td>
<td>0.344</td>
</tr>
<tr>
<td>4</td>
<td>74</td>
<td>68</td>
<td>6 [-3.3, 14.9]</td>
<td>0.375</td>
</tr>
<tr>
<td>Globally</td>
<td>75</td>
<td>64</td>
<td>11 [4.1, 17.7]</td>
<td>0.003*</td>
</tr>
</tbody>
</table>

WLI, white light imaging; CI, confidence interval. * Statistically significant.

tically significant 16.5% increase in the correct classification of images as true ulcerative or not (Table 1).

**True ulcerative image evaluation**
In the evaluation of 1000 true ulcerative images by all four investigators, they rated 493 and 715 images as real ulcerative with WLI-only and virtual chromoendoscopy-aided readings, respectively, and they rated 505 and 285 images as false ulcerative with WLI-only and virtual chromoendoscopy-aided readings, respectively. The improvement in the interpretation of these images as ulcerative or not was mirrored by a statistically significant 22% increase in accuracy (Table 1, Fig. 3).

For the 715 (of 1000) images correctly classified with virtual chromoendoscopy, the settings were rated as useful as follows: FICE 1 in 96%, FICE 2 in 92.2%, FICE 3 in 53.1%, and Blue mode in 66.3% of cases. On the other hand, for the 285 (of 1000) incorrectly classified images, the same settings were rated as useful (in fact misleading) as follows: FICE 1 in 53.3%, FICE 2 in 52.3%, FICE 3 in 49.5%, and Blue mode in 52.3% of cases.

**False ulcerative image evaluation**
For all four observers (and 200 false ulcerative images evaluated), in WLI 50 images were classified as true ulcerative and 150 as false ulcerative; on the other hand, when aided by virtual chromoendoscopy, the investigators classified 72 images as true ulcerative and 128 as false ulcerative, the second reading being associated with a statistically significant 11% decrease in the accuracy of correctly identifying the false ulcerative images (Table 1, Fig. 4).

For the 128 (of 200) correctly classified images, the virtual chromoendoscopy pre-sets were rated as useful as follows: FICE 1 in 50%, FICE 2 in 46.1%, FICE 3 in 39.1%, and Blue mode in 47.7% of cases. In contrast, for the 72 (of 200) incorrectly classified images, the same settings were rated as useful (in fact misleading) as follows: FICE 1 in 80.6%, FICE 2 in 79.2%, FICE 3 in 43.1%, and Blue mode in 75% of cases.

**Interobserver agreement**
The interobserver agreement was poor for both readings (WLI-only and virtual chromoendoscopy-aided), with a small decrease after virtual chromoendoscopy had been added (Table 2).

**Discussion**
VCE is an excellent examination technique for the small bowel, providing in the proper clinical context a diagnostic yield ranging from 55% to 81% [1, 10–14]. However, VCE is not a perfect examination tool. The entire surface of the small bowel mucosa is not visualized, at least not from the optimal or desired angles. Furthermore, bowel cleansing is suboptimal in a significant proportion of examinations, and the recognition and correct classification of mucosal lesions that are not clearly visualized remain challenging for the physician. Therefore, the development of technologies that can increase both the detection rates and diagnostic capabilities of VCE is desirable. Virtual chromoendoscopy with flexible spectral imaging (FICE) is based on narrowing the bandwidth of the three fundamental colors (red, green, and blue) that comprise normal WLI with dedicated post-acquisition processing software – so-called computerized spectral estimation technology [15]. On the other hand, Blue mode imaging uses a shift of the color coefficient in the short wavelength range of 490 to 430nm superimposed on regular WLI [2]. Both these technologies provide real-time enhancement of the surface patterns and color gradients of the gastrointestinal mucosa [2, 4], at least in theory helping the human eye to better perceive the small differences between adjacent mucosal areas. Recently, FICE technology and Blue mode have been incorporated into the RAPID software [2].
As expected, when FICE is applied in conventional endoscopy, the contrast of vascular and mucosal patterns is improved when it is applied in VCE. However, despite the presumed theoretical advantages, data on the application of virtual chromoendoscopy in VCE are very scarce. Only a few small studies [5 – 7, 16 – 21] and a limited number of case reports have been published to date [2, 3]. To summarize their findings, VCE with FICE improved the detectability of small-bowel lesions in the setting of opaque luminal content in a study of 12 VCE recordings [6], allowed the identification of a higher number of lesions (angioectasia and erosions) in a study of 20 patients [18], and enhanced the detection of angioectasia (but not of erosions, ulcerations, or tumors) in a study of 50 VCE recordings [5]; in contrast, in a study of 60 patients with occult gastrointestinal bleeding, FICE was not better than WLI in diagnosing significant lesions on small-bowel VCE other than in providing a more accurate characterization of some angioectasia [19], and in a study of 27 patients, Blue mode did not perform better than WLI in the calculation of the Lewis score [7]. Not only are the data contradictory and sparse; the studies also use different methodologies, so that the true ability of this technology to improve the rate of the detection of clinically significant lesions is called into question, and enthusiasm for virtual chromoendoscopy-aided VCE is tempered. However, a few reasons for the limited effectiveness of virtual chromoendoscopy-aided small-bowel VCE in previous studies can be stated. Clearly visible lesions in patients with excellent bowel preparation are very well visualized by WLI alone, and it is obvious that in these cases there will be no benefit of this technology. In keeping with this idea, a recent meta-analysis actually showed that small-bowel preparation with purgatives improves the diagnostic yield of the examination [22]. It should not come as a surprise, though, that in the presence of a prominent inflammatory involvement or clear mucosal lesions, virtual chromoendoscopy works well in improving the delineation of lesions, but it does not significantly improve their detection. In our view, it is the mucosal lesions that are difficult either to interpret or to visualize that need to be tested to demonstrate the gains expected from this technology. The superiority of the reconstructed virtual image over the standard endoscopic image in the evaluation of subtle lesions has not been assessed so far, ours being the first study to describe quantitatively the diagnostic performances of FICE and Blue mode settings in assisting the detection and characterization of difficult-to-interpret ulcerative lesions of the small-bowel mucosa. The study clearly demonstrates that virtual chromoendoscopy enhances the visibility of these lesions, making it easier for an experienced eye to recognize them. Indeed, by interfering with the reflectance spectrum in the gastrointestinal mucosa that has been altered by inflammation, virtual chromoendoscopy allows the identification of a higher number of erosions/ulcerations in comparison with conventional mode by highlighting inflammatory halos and increasing the contrast between pathological areas and surrounding mucosa. And although the ability of the evaluators to recognize lesions was quite low in WLI (49.4%), the use of virtual chromoendoscopy in our study improved accuracy in the correct characterization of erosions or ulcerations by as much as 22%, and only rarely (in 40 of 493 cases, or 8.1%) did virtual chromoendoscopy misguide the characterization of real ulcerative findings previously recognized in WLI as pathological. However, the false ulcerative lesions undoubtedly posed a challenge, inasmuch as the addition of virtual chromoendoscopy caused some of them (11%) to be erroneously misclassified as pathological, even by experienced readers; therefore care must be taken when virtual chromoendoscopy is used in the context of dirty luminal content. It is worth mentioning that at present there are no recommendations regarding which setting(s) to use for examining a potential mucosal lesion. One of the aims of the present study was to rate the usefulness of each setting, and our results show that the FICE 1 and 2 pre-sets are by far the most useful. Indeed, the results of various studies and expert opinions are concordant in that for the evaluation of ulcerative lesions of the small bowel, the FICE 1 and 2 pre-sets represent the wavelength combinations that generate the optimal difference between the spectral reflectance of normal mucosa and that of mucosa with pathological changes [2, 4, 6, 18, 19, 23]. The present study is not comparable with any of the previously mentioned studies. In this study, the lesion recognition rate was lower because of the inherent difficulty in image interpretation and the fact that the study design implied analyzing noncontinuous static images. It is therefore not surprising that when such difficult-to-interpret ulcerative lesions were considered, the kappa values were much lower (0.29 – 0.35 for WLI, compared with up to 0.66 in previous studies of more clearly visualized ulcerative lesions) [9, 23]. What is surprising in the present study is that the measures of agreement, although remaining within the same range of poor agreement, decreased when virtual chromoendoscopy was added (0.29 and 0.20 for true and false ulcerative lesions, respectively) despite an overall increase in better characterization of the pathological changes. This probably indicates that the reproducibility of the evaluation of difficult-to-interpret VCE images tends to decrease even more when the novel technology is added. Therefore, use of the FICE 1 and 2 pre-sets may be clinically meaningful in the characterization of difficult-to-interpret ulcerative mucosal small-bowel lesions. These pre-sets seem to provide a high level of visibility by increasing the transparency of opaque luminal content or by enhancing the color differences associated with ulcerated mucosa. However, the reader must be advised that the novel technology can also lead to the mistaken interpretation of artifacts as pathological features, and to avoid this, individual images should be evaluated only as part of a sequence in the recording. Extra caution is necessary in the setting of poor bowel preparation. Our proposed strategy would be first to evaluate the recording in WLI; then, if there is doubt about an ulcerative lesion, virtual chromoendoscopy FICE 1 and 2 pre-sets should be used to better define it. Although our results indicate the potential clinical usefulness of virtual chromoendoscopy-aided VCE, there are several limitations of this study. First of all, we do not know the proportion of false and real pathologic images encountered in the evaluation of VCE examinations, but it probably varies from patient to patient and from indication to indication. Moreover, in the quest for a real-life scenario, all the FICE settings were used together, and the ratings of their usefulness were subjective. The expert examiners evaluated only still images, not video recordings or successions of images, a fact that may have caused the role of virtual chromoendoscopy in image interpretation to be overestimated. Finally, whether the improvement presented herein really makes a difference in increasing the diagnostic accuracy of VCE remains unknown.

**Competing interests:** Cristiano Spada and Guido Costamagna are consultants for Given Imaging Ltd.
Acknowledgements

This work has been funded by the Executive Unit for Financing Higher Education, Research, Development and Innovation (UEFISCDI) through the contract PNII-IDEI 320/2007, and by the Sectoral Operational Programme Human Resources Development 2007–2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/1.5/S/132395.

References