Assessment of the Role of a Second Evaluation of Capsule Endoscopy Recordings to Improve Diagnostic Yield and Patient Management

Gerardo Blanco-Velasco a  Rolando Pinho b  Omar Michel Solórzano-Pineda a  Claudia Martínez-Camacho a  Luis Fernando García-Contreras a  Enrique Murcio-Pérez a  Oscar Victor Hernández-Mondragón a

a Endoscopy Service, Hospital de Especialidades, Centro Médico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Ciudad de México, Mexico; b Department of Gastroenterology, Centro Hospitalar Vila Nova de Gaia/Espinho, Vila Nova de Gaia, Portugal

Keywords
Capsule endoscopy · Small-bowel bleeding · Crohn’s disease

Abstract

Introduction: The diagnostic yield (DY) of small-bowel capsule endoscopy (SBCE) varies considerably according to its indication. Some strategies have been used to increase DY with varying results. The intention of this study was to identify whether evaluation of the SBCE recordings by a second reviewer can increase DY and change patient management.

Methods: One hundred SBCEs with different indications, already read by an endoscopist were read by a second blinded endoscopist. When the results of the 2 readings were different, the images were discussed by the endoscopists; if no consensus was reached, they took the opinion of a third endoscopist into account. All the participating endoscopists had experience in reading SBCEs (i.e., >50 per year). The SBCE findings were divided into positive (vascular lesions, ulcers, and tumors), equivocal (erosions or red spots), and negative. The interobserver agreement and the increase in DY were assessed as well as the percentage of false-negatives (FNs) in the first evaluation.

Results: The indications for SBCE were small-bowel bleeding (SBB) in 48 cases, Crohn’s disease (CD) in 30, and other causes (iron-deficiency anemia, small-bowel tumors, and diarrhea) in 22. There was substantial interobserver agreement between evaluations (κ = 0.79). The findings in the first evaluation were positive in 60%, equivocal in 20% and, negative in 20%. In the second evaluation, 66% were positive, 18% were equivocal, and 16% were negative. The increase in DY with the second reading was 6% (p = 0.380), i.e., 6.3% for SBB, 4.4% for CD and 9.2% for other indications, resulting in a change in management of 4% of the patients. FNs in the first SBCE reading were found in 4% of the SBCEs.

Discussion: A second evaluation of the SBCE recordings identified significant pathology that the first evaluation had missed, resulting in a nonsignificant 6% increase in DY and a change in the management of 4% of the patients.

© 2021 Sociedade Portuguesa de Gastrenterologia
Published by S. Karger AG, Basel

Palavras Chave
Endoscopia por cápsula · Hemorragia do intestino delgado · Doença de Crohn
Introduction

The diagnostic yield (DY) of small-bowel capsule endoscopy (SBCE) may vary depending on the indication. In obscure gastrointestinal bleeding (OGIB), the DY could be as high as 93% in ongoing overt bleeding and as low as 12.9% in previous overt bleeding [1]. Concerning other indications, the DY of SBCE is around 47% in iron-deficiency anemia (IDA) [2], 52% for Crohn’s disease (CD) [3], and 42.9% for diarrhea [4]. Some strategies, e.g., the use of simethicone and prokinetics, have been proposed to increase the diagnostic rate of capsule endoscopy with varying results [5].

There is limited evidence of the usefulness of a second evaluation of SBCE recordings. A study showed that a “back-to-back” evaluation increased the DY of the first from 37.5 to 62.5% [6]. “Back-to-back” in SBCE recordings has been used in other studies to assess DY but using different SBCE platforms [7, 8].

The little evidence that exists seems to show that the second evaluation can increase SBCE performance. The aim of this study is to assess whether a second evaluation of SBCE recordings by another reviewer can increase the DY and improve patient management.

Materials and Methods

Study Design

SBCE (Pillcam SB3; Given Imaging Ltd., Yoqneam, Israel) recordings with different indications, already read by a first endoscopist, were reread by a second blinded endoscopist. When the findings from the 2 revisions were different, the images were discussed by the endoscopists; if they did not reach consensus, they took the opinion of a third endoscopist into account. The first evaluation was considered as the finding of the first endoscopist only, while the second evaluation was considered as the finding of the second endoscopist plus the agreement reached in the case of different findings. All the participating endoscopists had experience in reading SBCE (i.e., >50 CE per year).

SBCE Selection

A total of 100 SBCEs with various indications, performed from August 2019 to October 2020, were included. All patients were prepared with 2 L of polyethylene glycol and an 8-h fast before the ingestion of the endoscopy capsule. Real-time review was used in all patients. In the event that the capsule did not pass into the duodenum after 60 min, metoclopramide (10 mg i.v.) was administered. If no progress was made after 90 min, the capsule was placed endoscopically in the duodenum. The belt and the data recorder were removed 12 h after they were placed.

Outcome Measures

The SBCE findings were classified by using capsule endoscopy structured terminology (CEST) [9] and divided in 3 groups: (1) positive, if significant lesions such as vascular lesions, ulcers, and tumors were found; (2) equivocal if nonspecific lesions such as erosions or red spots were found; and (3) negative, in the presence of irrelevant lesions or absence of findings. Only positive findings were considered as contributing to a positive DY. As per protocol, false-negative (FN) findings could only present in the first evaluation; they were defined as findings (either positive or equivocal) in the second evaluation not identified in the first evaluation. The recordings were read at a maximum speed of 10 frames per second in a single view, in line with the European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative [10].

Resumo

Introdução: O rendimento diagnóstico (DY) da enteroscopia por cápsula (SBCE) varia consideravelmente de acordo com as suas indicações. Foram estudadas algumas estratégias para melhorar o DY, com resultados variados. O objetivo deste estudo é avaliar se uma segunda leitura de registos de enteroscopia por cápsula por um segundo endoscopista pode melhorar o DY e alterar o manejo dos doentes.

Métodos: 100 SBCE realizadas por indicações variadas, previamente avaliadas por um endoscopista, foram reavalidadas por um segundo endoscopista cego para a primeira leitura. Na presença de resultados discordantes nas 2 leituras, as imagens foram discutidas pelos 2 endoscopistas, e por um terceiro na ausência de consenso entre os 2 primeiros. Todos os endoscopistas são experientes na leitura de SBCE (>50 SBCE/ano). Os achados foram categorizados em positivos (lesões vasculares, úlceras e tumores), equivocos (erosões e pontos vermelhos) e negativos. Foi avaliada a concordância inter-observador e o aumento do DY, bem como a percentagem de falsos negativos da primeira avaliação.

Resultados: As indicações para SBCE foram hemorragia do intestino delgado (SBB) em 48 casos, doença de Crohn (CD) em 30 e outras causas (anemia por défice de ferro, pesquisa de tumores do intestino delgado e diarreia) em 22. A concordância inter-observador foi substancial (k = 0.79). Os achados da primeira avaliação foram positivos em 60%, equivocos em 20% e negativos em 20%, enquanto na segunda avaliação foram positivos em 66%, equivocos em 18% e negativos em 16%. O aumento de DY foi de 6% com a segunda leitura (p = 0.38), sendo 6.3% para SBB, 4.4% para CD e 9.2% para outras indicações, resultando numa alteração no manejo de 4% dos doentes. Foram identificados falsos negativos na primeira avaliação em 4% das SBCE. Conclusão: A segunda avaliação dos registos de SBCE identificou patologia significativa previamente não identificada, resultando num aumento não significativo de 6% do DY e numa alteração no manejo de 4% dos doentes.
Sample Calculation

To calculate the sample size, the formula for the difference of 2 proportions was used. The calculation was based on the results of Min et al. [6] in which the DY of the first SBCE evaluation was 37.5% and the DY of the “back-to-back” SBCE was 62.5%. Epi Info software v3, considering a confidence interval (CI) of 95% and a statistical power of 80%, was used for sample size calculation. The estimated sample size was 70 patients per evaluation; however, to increase power, 100 SBCE recordings were included.

Statistical Analysis

Qualitative variables were calculated as frequencies and percentages. To study the relation between qualitative variables, the $\chi^2$ test was used. The interobserver agreement was assessed by using Cohen’s kappa ($\kappa$) coefficient. The $\kappa$ index was divided as slight agreement when the value was $<0.20$, fair at $0.21–0.40$, moderate at $0.41–0.60$, substantial at $0.61–0.80$, and almost perfect when $>0.81$. $p < 0.05$ was considered statistically significant. SPSS v22 (IBM, Chicago IL, USA) was used for statistical analysis.

Results

The indications for SBCE were small-bowel bleeding (SBB) in 48 cases, suspected or confirmed CD in 30, and other causes (IDA, small-bowel tumors, and diarrhea) in 22. There was good interobserver agreement between the 2 evaluations ($\kappa = 0.79$; Table 1).

Discussion

The interobserver agreement of SBCE findings has been reported as ranging from moderate to substantial ($\kappa = 0.48–0.71$) [11–14]. It also depends on the indications and the experience of the endoscopist. In cases of SBB, it can vary from $\kappa = 0.71$ for evaluations by senior endoscopists to $\kappa = 0.56$ for those made by junior endoscopists [13]. In CD, the interobserver agreement has been

---

**Table 1. Agreement between endoscopists by indications**

<table>
<thead>
<tr>
<th></th>
<th>SBB</th>
<th>CD</th>
<th>Other causes</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of cases</td>
<td>48</td>
<td>30</td>
<td>22</td>
<td>100</td>
</tr>
</tbody>
</table>

**Table 2. Findings in the first and second evaluations**

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th>Equivocal</th>
<th>Negative</th>
<th>$p$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>First evaluation</td>
<td>$60 (60)$</td>
<td>$20 (20)$</td>
<td>$20 (20)$</td>
<td></td>
</tr>
<tr>
<td>SBB</td>
<td>$32 (66.7)$</td>
<td>$9 (18.7)$</td>
<td>$7 (14.6)$</td>
<td></td>
</tr>
<tr>
<td>CD</td>
<td>$20 (66.7)$</td>
<td>$7 (23.3)$</td>
<td>$3 (10)$</td>
<td></td>
</tr>
<tr>
<td>Other causes</td>
<td>$8 (36.4)$</td>
<td>$4 (18.2)$</td>
<td>$10 (45.4)$</td>
<td>$0.66$</td>
</tr>
<tr>
<td>Second evaluation</td>
<td>$66 (66)$</td>
<td>$18 (18)$</td>
<td>$16 (16)$</td>
<td></td>
</tr>
<tr>
<td>SBB</td>
<td>$35 (73)$</td>
<td>$7 (14.5)$</td>
<td>$6 (12.5)$</td>
<td></td>
</tr>
<tr>
<td>CD</td>
<td>$21 (70)$</td>
<td>$8 (26.7)$</td>
<td>$1 (3.3)$</td>
<td></td>
</tr>
<tr>
<td>Other causes</td>
<td>$10 (45.5)$</td>
<td>$3 (13.6)$</td>
<td>$9 (40.9)$</td>
<td></td>
</tr>
</tbody>
</table>

---

In the first evaluation, findings were considered positive in 60%, equivocal in 20%, and negative in 20% of the cases. In the second evaluation, they were positive in 66%, equivocal in 18%, and negative in 16%. (Table 2). The changes in the type of findings from the second evaluation are detailed in Table 3. On 2 occasions, the intervention of a third endoscopist was required for consensus. These involved a gastrointestinal stromal tumor initially reported as normal and an ulcer diagnosed as “erosions and red spots” in the first evaluation.

The DY was 60% for the first evaluation and 66% after the second evaluation, resulting in a 6% increase ($p = 0.38$). The increased DY according to the indications was 6.3% for SBB, 4.4% for CD, and 9.2% for others. After the second evaluation, patient management was changed in 4 of these 6 patients; 1 of these 4 underwent surgery, 2 underwent therapeutic enteroscopy, and 1 had their medical therapy changed.

FNs were found in 4% of the SBCEs (cases 3, 4, and 6 in Table 3 and 1 case that went from normal to equivocal) after the second evaluation. The FN rate according to the indication was 2% for SBB, 6.7% for CD, and 4.5% for other indications.

---

**Table 3. Changes in positive findings after the second evaluation**

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Indication</th>
<th>First evaluation</th>
<th>Second evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IDA</td>
<td>erosions</td>
<td>ulcer</td>
</tr>
<tr>
<td>2</td>
<td>SBB</td>
<td>erosions + red spots</td>
<td>ulcer</td>
</tr>
<tr>
<td>3</td>
<td>SBB</td>
<td>normal</td>
<td>angioectasia</td>
</tr>
<tr>
<td>4</td>
<td>CD</td>
<td>normal</td>
<td>ulcer</td>
</tr>
<tr>
<td>5</td>
<td>SBB</td>
<td>erosions</td>
<td>angioectasia</td>
</tr>
<tr>
<td>6</td>
<td>IDA</td>
<td>normal</td>
<td>GIST</td>
</tr>
</tbody>
</table>

SBB, small-bowel bleeding; CD, Crohn’s disease; IDA, iron-deficiency anemia; GIST, gastrointestinal stromal tumor.

Values express $n$ (%). SBB, small-bowel bleeding; CD, Crohn’s disease.

---

In the first evaluation, findings were considered positive in 60%, equivocal in 20%, and negative in 20% of the cases. In the second evaluation, they were positive in 66%, equivocal in 18%, and negative in 16%. (Table 2). The changes in the type of findings from the second evaluation are detailed in Table 3. On 2 occasions, the intervention of a third endoscopist was required for consensus. These involved a gastrointestinal stromal tumor initially reported as normal and an ulcer diagnosed as “erosions and red spots” in the first evaluation.

The DY was 60% for the first evaluation and 66% after the second evaluation, resulting in a 6% increase ($p = 0.38$). The increased DY according to the indications was 6.3% for SBB, 4.4% for CD, and 9.2% for others. After the second evaluation, patient management was changed in 4 of these 6 patients; 1 of these 4 underwent surgery, 2 underwent therapeutic enteroscopy, and 1 had their medical therapy changed.

FNs were found in 4% of the SBCEs (cases 3, 4, and 6 in Table 3 and 1 case that went from normal to equivocal) after the second evaluation. The FN rate according to the indication was 2% for SBB, 6.7% for CD, and 4.5% for other indications.

**Discussion**

The interobserver agreement of SBCE findings has been reported as ranging from moderate to substantial ($\kappa = 0.48–0.71$) [11–14]. It also depends on the indications and the experience of the endoscopist. In cases of SBB, it can vary from $\kappa = 0.71$ for evaluations by senior endoscopists to $\kappa = 0.56$ for those made by junior endoscopists [13]. In CD, the interobserver agreement has been
reported as substantial ($\kappa = 0.68$) [15]. In our study, the
interobserver agreement was substantial and did not vary
according to the different indications.

The ESGE Quality Improvement Initiative mentions
that currently available data do not support a single optimal
DY per indication. For mixed indications, the DY varies
between 27 and 77.3% and, for suspected gastrointestinal
bleeding, between 31 and 68% [10]. In this study, the DY
in the first evaluation was already high at 60%, ranging from
66.7% for SBB and CD to 36.4% for other indications.

There have been several attempts to improve the DY
of SBCE. The use of small-bowel preparation, antifoam-
ing agents, or prokinetics has been proposed. Recent sys-
tematic reviews and meta-analyses to evaluate whether
bowel preparation before SBCE improves the DY con-
cluded that there are no clear advantages [16–18]. It has
been suggested that a “second-look” with another SBCE
can increase the DY. Svarta et al. [19] found positive find-
ings in 55% of repeated SBCEs that resulted in a change
in management in 39% of the patients. In our study, a
second evaluation of the same capsule increased the DY
by 6%; however, it was time-consuming and no statisti-
cally significant difference was found.

Van de Bruaene et al. [20] identified FNs in 9% of the
SBCEs. They defined FNs as bleeding sources located in
the small bowel but not diagnosed during the initial
SBCE. In our study, FNs were found in 4%, but our defi-
nition of FN was different.

The final diagnosis was changed in 6 cases. Three of
these were initially diagnosed as erosions and red spots; in
the other 3, the SBCE results were reported as normal. Al-
though not reaching statistical significance, these findings
are nonetheless of obvious relevance. In 4% of the patients,
a change in management resulted from the reevaluation. In
1 case, a gastrointestinal stromal tumor not diagnosed at the
first evaluation and requiring surgical intervention was
found. Another 2 patients required therapeutic endoscopy
using argon plasma coagulation to treat angioectasia. One
patient diagnosed as having CD had a change in medical
therapy after ulcers were found at the second evaluation.
The diagnosis remained nonsteroidal anti-inflammatory
drug (NSAID)-induced enteropathy in 2 patients.

The lack of statistical significance is likely due to the
high DY present in the first evaluation. It is at the upper
end of previously reported data and similar to the DY ob-
tained after the “back-to-back” approach in the study by
Min et al. [6] used to calculate the sample size. This high
DY probably reflects improved patient selection and im-
provements in SBCE technology in recent series. Further
studies evaluating gains in DY after interventions, like
second evaluations, the use of bowel preparations, or new
protocols, should take into account the higher initial DY
and marginal increases in DY.

**Conclusion**

Relevant findings missed in a first evaluation are often
identified at a second evaluation, resulting in relevant
changes in patient management. Due to the time-con-
suming nature of SBCE reading, a second evaluation
could be selectively offered to negative SBCE patients
with suspected pathology. Further studies evaluating in-
terventions to improve the DY of SBCE should take into
account the higher initial DY than was previously report-
ed, and, consequently, more modest gains in DY.

**Acknowledgment**

We did not have assistance with this article.

**Statement of Ethics**

This study is in line with the regulations of the Mexican General
Health Law regarding health research and the Declaration of Hel-
sinki as well as the rules and institutional instructions on scientific
research. SBCE is proven to be safe, so this was classified as a without-
risk study. The study was performed under authorization of the hos-
pital’s Ethics and Research Committee (R-2020-3601-161).

**Conflict of Interest Statement**

G.B.-V. has participated in speaking activities for Medtronic
and Jinshan Science and Technology Co. The other authors have
no conflicts of interest to declare.

**Funding Sources**

We did not receive financial assistance for this article.

**Author Contributions**

G.B.-V. made substantial contributions to the conception, acquisi-
tion, analysis, and interpretation of data for the work. R.P. made
substantial contributions to the conception, analysis, and interpreta-
tion of data for the work. O.M.S.-P., C.M.-C., and L.F.G.-C. made
substantial contributions to the conception and acquisition of data
for the work. E.M.-P. agreed to be accountable for all aspects of the
work in ensuring that questions related to the accuracy or integrity
of any part of the work are appropriately investigated and resolved.
O.V.H.-M. gave final approval of the version to be published.
References


