Thieme

More than 20 procedures are necessary to learn small bowel capsule endoscopy: Learning curve pilot study of 535 trainee cases

OPEN ACCESS

Authors

Anders Bo Nielsen[1,2,3,4], Michael Dam Jensen[5,6], Jacob Broder Brodersen[5,6], Jens Kjeldsen[2,7], Christian B. Laursen[8,9], Lars Konge[10,1], Stig Borbjerg Laursen[2,7]

Institutions

1 Research Unit of Medical Education, Department of Clinical Research, University of Southern Denmark, Odense, Denmark
2 Department of Medical Gastroenterology, Odense University Hospital, Odense, Denmark
3 Research Unit of OPEN - Open Patient data Explorative Network , University of Southern Denmark Department of Clinical Research, Odense, Denmark
4 Department of Anesthesiology and Intensive Care, Svendborg, Odense University Hospital, Odense, Denmark
5 Department of Regional Health Research, University of Southern Denmark, Odense, Denmark
6 Internal Medicine, Section of Gastroenterology, Hospital South West Jutland, Esbjerg, Denmark
7 University of Southern Denmark Department of Clinical Research, Odense, Denmark
8 Department of Respiratory Medicine, Odense University Hospital, Odense, Denmark
9 Odense Respiratory Research Unit (ODIN), University of Southern Denmark Department of Clinical Research, Odense, Denmark
10 Copenhagen Academy for Medical Education and Simulation, Capital Region of Denmark, Copenhagen, Denmark

Keywords

Endoscopy Small Bowel, Capsule endoscopy, Quality and logistical aspects, Training, Quality management

received 27.10.2023
accepted after revision 13.3.2024

Bibliography

Endosc Int Open 2024; 12: E697–E703
DOI 10.1055/a-2308-1613
ISSN 2364-3722

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Corresponding author

Dr. Anders Bo Nielsen, MD, Odense University Hospital, J. B. Winsløws Vej 4, 5000 Odense, Denmark
anders.bo.nielsen@rsyd.dk
andersbonielsen@hotmail.com

ABSTRACT

Background and study aims The number of procedures needed to acquire a sufficient level of skills to perform an unassisted evaluation of small bowel capsule endoscopy (SBCE) is unknown. We aimed to establish learning curves, diagnostic accuracy, and the number of procedures needed for reviewing small bowel capsule endoscopies unassisted.

Methods An expert panel developed a 1-day course including lessons (examination, anatomy, and pathology) and hands-on training. After completing the course, participants received 50 cases in a randomized sequence. An interactive questionnaire about landmarks, findings, and diagnosis followed each case. After submitting the questionnaire, participants received feedback. Data are presented using CUSUM (cumulative sum control chart) learning curves and sensitivity/specificity analyses compared with expert opinions.

Results We included 22 gastroenterologists from 11 different Danish hospitals. A total of 535 cases were reviewed (mean: 28; range: 11–50). CUSUM plots demonstrated learning progression for diagnosis and findings during the course, but none of the participants reached a learning plateau with sufficient competencies. The sensitivity for all findings was 65% (95% confidence interval [CI] 0.51–0.82) for the first 20 procedures and 67% (95% CI 0.58–0.73) from case 21 until completion or dropout. The specificity was 63% (95% CI 0.52–0.74) for the first 20 procedures and 57% (95% CI 0.37–0.77) for the rest.

Conclusions Our data indicate that learning SBCE may be more difficult than previously recognized due to low discriminative abilities after 20 cases except for the identification of CD. This indicates that 20 SBCE cases may not be sufficient to achieve competency for reviewing SBCE without supervision.

Introduction

Capsule endoscopy (CE) is a minimally invasive method for imaging the bowel mucosa introduced well over 20 years ago [1, 2, 3]. Small bowel capsule endoscopy (SBCE) plays a major role in evaluation of patients with Crohn's disease (CD), iron deficiency anemia, and obscure gastrointestinal bleeding [4, 5, 6].

CE is an effective modality for diagnosing diseases in the small bowel, but it is sensitive to the experience and level of skills of the reviewer, and it is known that interobserver variation can occur even among experienced gastroenterologists [7]. Studies have shown a significant difference in the accuracy of CE reviewers depending on experience level; thus, it may be beneficial for future CE reviewers to participate in a structured training program [8, 9].

Both national and international societies recommend training before performing reviews of SBCE [10, 11]. The latest statement by the European Society for Gastrointestinal Endoscopy (ESGE) indicates that learning SBCE requires a minimum of 75 to 100 SBCEs/year for a center, experience with bidirectional endoscopy is desirable for structured SBCE training, 50% hands-on training at SBCE courses, and competencies in SBCE evaluation can be assessed after 30 procedures with direct observation, short videos, or multiple-choice questionnaires [12]. Similar key components are recommended by The American Society for Gastrointestinal Endoscopy (ASGE) with 20 supervised SBCE-procedures before practicing independently [11].

However, there is no structured evidence-based learning program based on these recommendations and the number of supervised cases needed remains unknown. Few studies have investigated different learning models, but common for them was the evaluation of a low number of SBCE cases, small sample sizes, and inconclusive results [8, 13, 14].

The current study aimed to establish learning curves for SBCE trainees, study the diagnostic accuracy of SBCE trainees, and finally assess the number of SBCE procedures needed to learn SBCE.

Materials and methods

Setting

The study was designed as a prospective learning study at Odense University Hospital, Denmark. All cases were captured with SB2 or SB3 PillCam and the Rapid PillCam Reader v9 software (Medtronic, Minnesota, United States) was used. The study was reported according to the Standards for QUality Improvement Reporting Excellence in Education (SQUIRE-EDU) [15]. All participants were informed about the study and provided informed written consent.

Content development

A panel of three experts (JK, professor; MDJ, consultant; JBB, specialist) in gastroenterology and SBCE developed an educational program. Consensus was reached on a 1-day course including lectures and hands-on training as recommended by ESGE [12].

The course consisted of four lectures on the following topics: 1) the technology and clinical use of SBCE (40 minutes); 2) evaluation of SBCE (40 minutes); 3) normal findings and anatomical landmarks (50 minutes); and 4) most common pathologies in the small intestine (60 minutes). Two hands-on modules including software exercises with three normal cases and two cases with pathology and a plenum evaluation (120 minutes) also were part of the course.

The experts designed a series of 50 cases based on anonymized real-life SBCEs. The cases were organized with a medical history, a corresponding unedited SBCE video, and an interactive questionnaire. The interactive questionnaire was designed to give feedback and corrections to trainees. The answers to the questionnaire were used to monitor skills development. Currently, no one has identified the optimal distribution of cases for learning evaluation of SBCE, so this study aimed to follow the statement by ESGE [12]. We included SBCE cases with CD (n = 23), cases without pathological findings categorized as normal (n = 12), bleeding (n = 10), tumors (n = 4), and stenosis (n = 1) [16]. Lists with correct answers for findings and diagnoses of the 50 SBCEs were developed for each SBCE by the experts. In case of disagreement between most of the participants and the expert assessment, a second review was performed by the experts. Cases were renamed to ensure blinding.

Participants

Physicians were at least second-year residents in gastroenterology with experience in both upper and lower endoscopy. Prior experience with SBCE was an exclusion criterion.

Data collection

Data were collected in the online database, Research Electronic Data Capture (REDCap), using online questionnaires [17, 18]. Each case had the same matching questionnaire with a standard range of multiple-choice questions and landmark recognition tasks. Participants received both written and oral instructions on the course and cases on the course day along with teaching resources. All participants received an individually randomized sequence of cases. Participants who were inactive for ≥ 3 months were excluded.

Outcome measurements

The primary outcome was the percentage of total number of correct evaluated cases based on correct identification of pathology/lesions based on a multiple-choice questionnaire including nine options – erosions, ulcers, angioectasia, polyps, tumors, stenosis, lymphangiectasia, bleeding, or no pathological findings – and using the expert consensus agreement of each case as the reference test. Multiple selections were accepted.

Second, the questionnaire asked for a diagnosis based on the findings, which included the following diagnoses: normal (no pathology), CD, small bowel bleeding, small bowel tumor, or other. Participants were also asked to identify landmarks with the indication of the time for passing the gastroesophageal junction, pylorus, and ileocecal junction. A correct landmark identification was defined as a time indication within 30 seconds of the time stated in the list with correct answers by the experts. Finally, self-reported time consumption was noted in the questionnaire.

Statistical analysis

Statistical analysis was carried out in SPSS statistics version 28 (IBM, New York, United States).

To evaluate the diagnostic accuracy of the participants, we analyzed the sensitivity and specificity of the diagnosis and the specific findings. 95% confidence intervals were calculated. The results were analyzed using Fisher's exact test. Two-tailed P <0.05 was considered statistically significant.

Cumulated sum analysis was used to calculate the learning curves of participant abilities in reviewing SBCE. Calculations in our study were based on Bolsin and Colson's explanations [19]. Acceptable (p_0 = 0.1) and unacceptable (p_1 = 0.2) failures were defined to calculate the value of a successful review (s) and the penalty for an unsuccessful review (1-s).

$$s = \ln((1-p_1)/(1-p_0))/(\ln((1-p_1)/(1-p_0)) + \ln((1-p_0)/(1-p_1)))$$

The acceptable failure rate (p_0) was designated as 10% and 20% for the unacceptable (p_1). This resulted in a value of successful review(s) at 0.15 and a penalty of 0.85.

A cumulative sum (CUSUM) score and graph were based on these values and said to signal when the predefined decision interval (H) was crossed. A learning curve was established, and the decision intervals were repeated and stacked graphically as horizontal lines to determine when a learning plateau of competencies was acquired. We used an α and β of 0.1 to produce an easily interpretable graph as the acceptable and unacceptable performance decision intervals were equal. The predetermined decision interval H can be divided into an interval (H_1) between the acceptable levels and between the unacceptable levels (H_0), both were calculated to be 2.71.

They were calculated as:

$$H_1 = (\ln((1-\beta)/\alpha)/(\ln((1-p_1)/(1-p_0)) + (\ln((1-p_0)/(1-p_1)))$$

$$H_0 = (\ln((1-\alpha)/\beta)/(\ln((1-p_1)/(1-p_0)) + (\ln((1-p_0)/(1-p_1)))$$

The degree of the slope at the CUSUM curve is a measure of the learning progress in mastering the evaluation of SBCE. An upward deflection of the curve is a result of slow learning and a low level of skills in mastering the procedure, while a flattened horizontal line is a sign of mastering skills. This might be followed by a downward deflection of the line, which also indicates mastering the skill. The greater the slope, the slower the learning progress [20].

We aimed to include 20 participants based on the sparse knowledge within the field [8]. Because no previous studies used similar outcomes, there was no satisfactory basis for a power calculation to test current recommendations for learning SBCE.

▶ **Table 1** Participant demographics and characteristics.

	Participants
Total, n	22
Female, %	45
Mean age, years (range)	37 (29–54)
Years as a doctor (range)	9 (2–25)
Years in gastroenterology (range)	7 (1–20)
Number of specialists, n	4
Number of registrars, n	15

Results

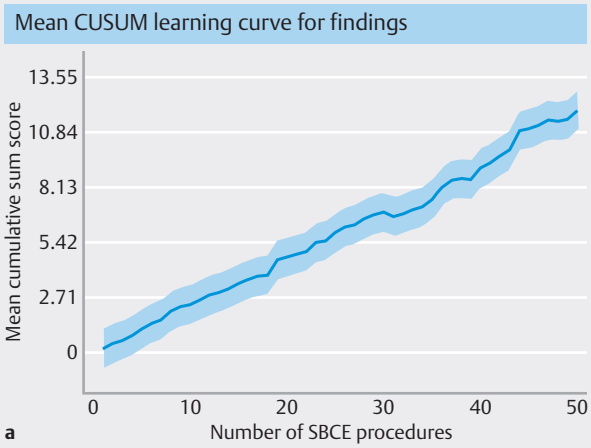
Eighteen registrars and four specialists (n = 22) in gastroenterology from Danish hospitals were included in the study (▶ **Table 1**). The mandatory 1-day course was held in October 2018 at Odense University Hospital, Denmark and the online case program was open for answers from October 2018 until February 2020. A total of 535 cases were reviewed with a mean of 28 cases (range: 11–50). Seventeen participants completed at least 15 cases, 10 completed 20 cases, seven completed 30 cases, and four all 50 cases. Three of the registrars did not review any cases.

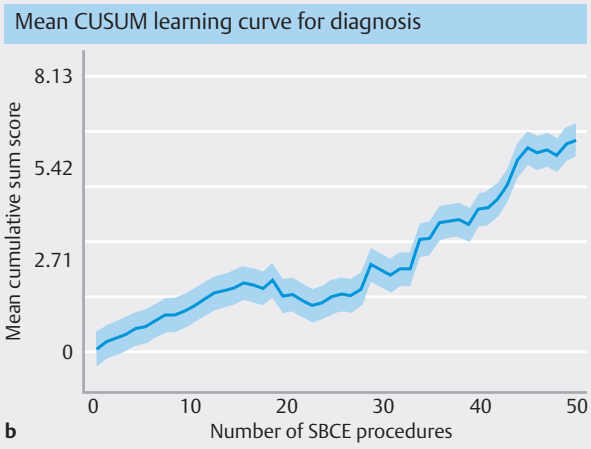
CUSUM

Results are presented as summarized learning curves for both findings and diagnosis in ▶ **Fig. 1**. The graphical data demonstrate that the participants did not achieve sufficient competency during the entire study period and that none of the participants reached a persistent learning plateau in both identifications of findings (Fig. 1a) and establishing the correct diagnosis (▶ **Fig. 1b**).

Discriminative abilities

The mean sensitivity for all findings was 65% (95% confidence interval [CI] 0.51–0.82) for the first 20 procedures and 67% (95% CI 0.58–0.73) from Case 21 until completion or dropout. Regarding specific findings, the sensitivity for angioectasias was best at 80% (95% CI 0.65–0.95), 79% (95% CI 0.75–0.83) for erosions, and 72% (95% CI 0.52–0.92) for ulcers. For the first 20 cases, the sensitivities for each finding were 71% (95% CI 0.56–0.86) for angioectasias, 71% (95% CI 0.44–0.98) for ulcers, and 79% (95% CI 0.73–0.85) for erosions. In comparison,





▶ **Fig. 1** CUSUM graphs. **a** CUSUM plot illustrating the mean score of findings for the participants. 95% confidence intervals (CIs) are indicated by the dashed lines. **b** CUSUM plot illustrating mean score of diagnosis for the participants. 95% CIs are indicated by the dashed lines.

the sensitivities for Case 21 until completion/dropout were 85% (95% CI 0.65–0.95) for angioectasias, 74% (95% CI 0.49–0.99) for ulcers, and 79% (95% CI 0.75–0.83) for erosions.

There was no apparent difference in the sensitivities for determining the correct diagnosis (76%; 95% CI 0.65–0.85) when comparing the first 20 cases (76%; 95% CI 0.55–0.92) and the last completed cases (76%; 95% CI 0.58–0.85). The sensitivity for CD was highest at 89% (95% CI 0.84–0.90). However, the sensitivity for CD was unchanged between the first 20 cases and the rest of the cases (89% [95% CI 0.83–0.95] vs 89% [95% CI 0.84–0.94]). There was also no difference in sensitivity for identifying small bowel bleeding between the first 20 cases (74%; 95% CI 0.64–0.84), the last cases (72%; 95% CI 0.52–0.92), and overall (73%; 95% CI 0.65–0.83). All cases with capsule retention were identified.

The overall specificity for pathological findings was 46% (95% CI 0.36–0.56). There were no changes in specificity from Cases 1 to 20 to the rest of the cases. The specificity for a correct diagnosis was 62% (95% CI 0.52–0.72); 63% (95% CI 0.52–0.74) for the first 20 procedures, and 57% (95% CI 0.37–0.77) for the last cases. 37% of the cases categorized as normal were mistaken for CD. Results for all findings are shown in ▶ **Table 2**. ▶ **Table 3** shows the rate of correct diagnosis for each SBCE case.

The mean rate of total correctly identified landmarks of passages for Cases 1 to 20 was 66% (95% CI 0.63–0.69), and 70% (95% CI 0.67–0.73) after Case 20. There was a significant improvement between Cases 1 to 20 and after Case 20 for recognition of passage to pylorus ($P = 0.029$), while no significant difference was found between Cases 1 to 20 and after Case 20 for the other landmarks and the mean rate of all landmarks in total.

Four participants achieved a sensitivity higher than 90% in recognizing CD, two participants in recognizing tumors, and one participant in recognizing the examination as normal after

▶ **Table 2** Sensitivity and specificity analysis for each finding.

Findings	Total number (n)		Sensitivity (%)			Specificity (no findings) (%)		
	Obtainable	Identified	1–20	21–	Total	1–20	21–	Total
Erosions	280	222	79 (73–85)	79 (75–83)	79 (75–83)	58 (52–64)	59 (49–69)	59 (55–63)
Ulcers	115	82	71 (44–98)	74 (49–99)	72 (52–92)	64 (60–68)	66 (56–76)	64 (60–68)
Angioectasias	66	53	71 (56–86)	85 (74–94)	80 (65–95)	63 (57–69)	66 (56–76)	64 (59–69)
Polyps	24	5	11 (–)	60 (–)	22 (–)	66 (56–76)	67 (56–78)	67 (62–72)
Tumors	23	11	46 (–)	50 (–)	48 (35–61)	65 (58–72)	68 (59–77)	66 (60–72)
Stenosis	61	33	46 (31–61)	50 (41–59)	50 (41–59)	65 (57–73)	69 (57–81)	66 (61–71)
Lymphangiectasias	39	14	30 (6–54)	46 (26–66)	37 (17–57)	66 (59–73)	68 (58–78)	67 (62–72)
Bleedings	115	80	63 (51–75)	71 (62–80)	70 (61–79)	64 (57–71)	66 (53–79)	65 (59–71)
No pathological findings	125	57	69 (61–77)	71 (62–80)	69 (65–73)	48 (27–61)	41 (30–52)	46 (36–56)
Total	723	500	65 (51–79)	67 (58–76)	65 (58–72)	65 (58–72)	67 (62–72)	65 (61–69)

▶ **Table 3** Cases, diagnosis, and rate of correct diagnosis among trainees.

Case number	n, completed	Diagnosis	Rate of correct diagnosis among trainees
1	10	CD	90
2	10	Bleeding	50
3	8	CD	100
4	15	CD	80
5	10	Normal	78
6	7	Normal	29
7	13	Bleeding	8
8	11	CD	82
9	11	Bleeding	100
10	13	CD	92
11	12	CD	83
12	12	Bleeding	58
13	10	Tumor	40
14	10	CD	100
15	9	CD	89
16	10	CD	90
17	9	Normal	67
18	9	Normal	78
19	11	Bleeding	82
20	12	Tumor	42
21	10	CD	80
22	10	Normal	90
23	11	Normal	64
24	12	Tumor	42
25	9	CD	100
26	10	CD	10
27	13	Normal	69
28	11	Tumor	55
29	11	Normal	27
30	13	Normal	54
31	12	Bleeding	75
32	9	CD	78
33	9	CD	100
34	12	Normal	42
35	13	Bleeding	92
36	12	Bleeding	75
37	9	Normal	33
38	12	Normal	25
39	12	CD	100
40	13	CD	85
41	11	CD	73
42	9	CD	100
43	9	Bleeding	100
44	11	Bleeding	91
45	10	CD	90
46	10	CD	100
47	10	Other	100
48	10	CD	90
49	8	CD	100
50	12	CD	100
Total	535		74

CD, Crohn's disease.

completing the first 20 training cases. None of these participants had a sensitivity > 90% for more than one diagnosis.

Time used for evaluation

The mean time for evaluation of a SBCE was 42.2 minutes (95% CI 33.2–51.2). The mean time used for Cases 1 to 5 was 58.2 minutes (95% CI 48.2–68.2), Cases 10 to 15 38.4 minutes (95% CI 33.4–43.4), Cases 20 to 25 44.9 minutes (95% CI 38.9–50.9), Cases 30 to 35 37.6 minutes (95% CI 34.6–40.6), and Cases 45 to 50 34.1 minutes (95% CI 32.1–36.1). There was a significant decrease in the time used between Cases 1 to 5 and 10 to 15 ($P = 0.028$), and between Cases 20 to 25 and 45 to 50 ($P = 0.006$).

Expert reevaluations

Cases 7 (bleeding), 26 (CD), and 38 (normal) were selected for a second review due to discrepancies in the answers from the participants and the list of correct answers by the experts. All experts agreed on the findings in all three cases and there was full agreement on the diagnosis.

Discussion

The present study evaluated subsequent development of reviewing skills in SBCE by establishing learning curves, diagnostic accuracy, and the number of procedures needed to learn SBCE.

Learning curves

Mean CUSUM scores for the learning curve for SBCE-diagnosis (▶ **Fig. 1b**) leveled out after completing 15 cases, which reflects a learning plateau and attainment of some competencies in line with previous studies describing learning curves for SBCE [8,

21, 22] and the latest position statement by ESGE and the learning curriculum suggested by ASGE [11, 12]. In our study, this plateau ends after 28 reviews as the learning curve takes another step upward until the 50 cases are completed. There was no sign of achieving competencies in the last 22 cases because the learning curve did not level out or begin to decrease. The learning curve for findings was nearly linear during the 50 cases without any plateau or flattening, which indicates that the participants still were in a significant learning phase with a high failure rate. The learning curve for making the correct diagnosis was linear until 18 completed cases.

The relentless rise during the first 50 procedures indicates that the participants still were in a learning phase. This was also supported by the lack of improvement in sensitivity and specificity rate from the first 20 cases to the completed cases after number 20, which also indicates the absence of achieving sufficient competencies.

These findings call into question the previous recommendations because the participants did not completely attain competencies and sufficient ability to identify the right findings, which may lead to error in patient diagnosis and treatment [11, 12].

Discriminative abilities

We found no significant improvement in participant ability to identify specific findings or identify the right diagnosis between the first 20 cases and the last completed cases. Despite completing 20 SBCE cases, the observed sensitivities for ulcers, polyps, tumors, stenoses, and small bowel bleeding continued to be low (50%–74%). Likewise, despite completing 20 previous cases, only 57% of normal cases were classified correctly. This underscores the difficulty in diagnosing SBCE without pathology. It is well known that intestinal debris can be mistaken as ulcerations, but without edema and surrounding redness, which might be the explanation for the normal cases classified as CD [23].

Regarding diagnosis, only CD had a relatively high sensitivity of 89%. Corresponding to real-life patients, these findings are thought-provoking because treatments often are based on the evaluation and SBCE is one of the reference standards for excluding small bowel disorders [22]. Nevertheless, the missing improvement between the first 20 cases and the following in identifying the anatomical landmarks were the same low rates as for findings and diagnosis. We only experienced a significant improvement for the duodenal landmark, and that improvement was not impressive with a correct rate of 68%, which is too low to have a clinical impact as recommended by ESGE [12]. On the other hand, the participants demonstrated good competencies in all cases in recognizing the landmark of passage to the stomach in more than 90% of the cases.

We showed a clear decrease in time consumption for reviewing SBCE cases throughout the study. This can be seen as a sign of missing dedication and prioritization during a busy workday, too high reading speed, or a sign of confidence without achieving enough competencies, which is alarming [24]. The rapidly developing use of artificial intelligence (AI) with promising diagnostic accuracy can potentially assist learning and change the reviewing process toward deep learning algorithms instead of in-person evaluation [25]. The use of AI to assist in learning SBCE calls for further studies to ensure sufficient learning and diagnostic accuracy by the reviewers.

Strength and weaknesses

Our study is the first to explore the effects of a structured course followed by 50 randomized training cases and is the largest based on completed cases and with the longest prospective follow-up period for learning SBCE. Another strength is the use of a web-based platform to deliver feedback, corrections, and new cases because it allows for blinded and objective evaluation of the participants, which can be difficult when training in one's own department with colleagues or supervisors [26, 27]. Moreover, all participants in this study had a relevant specialty and educational status to ensure their ability and readiness to learn and achieve new competencies in evaluation of SBCE.

Supervision was limited to lists with correct answers, illustrative pictures of findings, and answer corrections. Participants were not able to discuss findings in person with an experienced SBCE reviewer. Another limitation was the lack of participants completing all 50 cases (n = 4), and the fact that only seven participants completed more than 30 cases as recommended by ESGE, but without reaching a learning plateau. This can be seen as a commitment challenge and may be due to the amount of time used to complete a case when you are a novice, which might be addressed by exposing participants to shorter video sequences.

Conclusions

The present pilot study indicates that learning SBCE may be more difficult than previously recognized and that trainees who have completed 20 procedures continue to have low discriminative abilities except for the identification of CD. Our findings indicate that more than 20 supervised procedures are needed to achieve sufficient competencies for assessing SBCE without supervision; however, this requires further exploration.

Conflict of Interest

The authors declare that they have no conflict of interest.

Clinical trial

ClinicalTrials.gov (http://www.clinicaltrials.gov/)
Registration number (trial ID): NCT04186390
Type of Study: Prospective

References

[1] Eliakim R, Fireman Z, Gralnek IM et al. Evaluation of the PillCam Colon capsule in the detection of colonic pathology: results of the first

multicenter, prospective, comparative study. Endoscopy 2006; 38: 963–970 doi:10.1055/s-2006-944832

[2] Eliakim R, Yassin K, Niv Y et al. Prospective multicenter performance evaluation of the second-generation colon capsule compared with colonoscopy. Endoscopy 2009; 41: 1026–1031

[3] Kakugawa Y, Saito Y, Saito S et al. New reduced volume preparation regimen in colon capsule endoscopy. World J Gastroenterol 2012; 18: 2092–2098 doi:10.3748/wjg.v18.i17.2092

[4] Gerson LB, Fidler JL, Cave DR et al. ACG Clinical Guideline: Diagnosis and Management of Small Bowel Bleeding. Am J Gastroenterol 2015; 110: 1265–1287

[5] Goddard AF, James MW, McIntyre AS et al. Guidelines for the management of iron deficiency anaemia. Gut 2011; 60: 1309–1316 doi:10.1136/gut.2010.228874

[6] Raju GS, Gerson L, Das A et al. American Gastroenterological Association (AGA) Institute technical review on obscure gastrointestinal bleeding. Gastroenterology 2007; 133: 1697–1717 doi:10.1053/j.gastro.2007.06.007

[7] Lai LH, Wong GL, Chow DK et al. Inter-observer variations on interpretation of capsule endoscopies. Eur J Gastroenterol Hepatol 2006; 18: 283–286 doi:10.1097/00042737-200603000-00009

[8] Lim YJ, Joo YS, Jung DY et al. Learning curve of capsule endoscopy. Clin Endosc 2013; 46: 633–636 doi:10.5946/ce.2013.46.6.633

[9] Koffas A, Laskaratos FM, Epstein O. Training in video capsule endoscopy: Current status and unmet needs. World J Gastrointest Endosc 2019; 11: 395–402 doi:10.4253/wjge.v11.i6.395

[10] Tai FWD, Parker C, Sidhu R et al. Training pathway for small bowel capsule endoscopy in the UK. Frontline Gastroenterol 2021; 13: 206–210 doi:10.1136/flgastro-2021-101810

[11] Rajan EA, Pais SA, Degregorio BT et al. Small-bowel endoscopy core curriculum. Gastrointest Endosc 2013; 77: 1–6 doi:10.1016/j.gie.2012.09.023

[12] Sidhu R, Chetcuti Zammit S, Baltes P et al. Curriculum for small-bowel capsule endoscopy and device-assisted enteroscopy training in Europe: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. Endoscopy 2020; 52: 669–686

[13] Postgate A, Haycock A, Fitzpatrick A et al. How should we train capsule endoscopy? A pilot study of performance changes during a structured capsule endoscopy training program Dig Dis Sci 2009; 54: 1672–1679

[14] Sidhu R, Sakellariou P, McAlindon ME et al. Is formal training necessary for capsule endoscopy? The largest gastroenterology trainee study with controls Dig Liver Dis 2008; 40: 298–302

[15] Ogrinc G, Armstrong GE, Dolansky MA et al. SQUIRE-EDU (Standards for QUality Improvement Reporting Excellence in Education): Publication Guidelines for Educational Improvement. Acad Med 2019; 94: 1461–1470 doi:10.1097/ACM.0000000000002750

[16] Kharazmi AA, Aslani S, Kristiansen MF et al. Indications and diagnostic yield of small-bowel capsule endoscopy in a real-world setting. BMC Gastroenterol 2020; 20: 177 doi:10.1186/s12876-020-01326-8

[17] Harris PA, Taylor R, Thielke R et al. Research electronic data capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009; 42: 377–381 doi:10.1016/j.jbi.2008.08.010

[18] Harris PA, Taylor R, Minor BL et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019; 95: 103208 doi:10.1016/j.jbi.2019.103208

[19] Bolsin S, Colson M. The use of the Cusum technique in the assessment of trainee competence in new procedures. Int J Qual Health Care 2000; 12: 433–438 doi:10.1093/intqhc/12.5.433

[20] Yap CH, Colson ME, Watters DA. Cumulative sum techniques for surgeons: a brief review. ANZ J Surg 2007; 77: 583–586 doi:10.1111/j.1445-2197.2007.04155.x

[21] Albert JG, Humbla O, McAlindon ME et al. A simple evaluation tool (ET-CET) indicates increase of diagnostic skills from small bowel capsule endoscopy training courses: A prospective observational european multicenter study. Medicine 2015; 94: e1941

[22] Rondonotti E, Soncini M, Girelli CM et al. Can we improve the detection rate and interobserver agreement in capsule endoscopy? Dig Liver Dis 2012; 44: 1006–1011 doi:10.1016/j.dld.2012.06.014

[23] Girelli CM, Porta P, Malacrida V et al. Clinical outcome of patients examined by capsule endoscopy for suspected small bowel Crohn's disease. Dig Liver Dis 2007; 39: 148–154 doi:10.1016/j.dld.2006.10.018

[24] Spada C, McNamara D, Despott EJ et al. Performance measures for small-bowel endoscopy: A European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. United European Gastroenterol J 2019; 7: 614–641 doi:10.1177/2050640619850365

[25] Mascarenhas M, Mendes F, Ribeiro T et al. Deep learning and minimally invasive endoscopy: automatic classification of pleomorphic gastric lesions in capsule endoscopy. Clin Transl Gastroenterol 2023; 14: e00609 doi:10.14309/ctg.0000000000000609

[26] Downing SM YR, Haladyne TM, Axelson RD et al. Assessment in Health Professions Education. 2009

[27] Kilminster S, Cottrell D, Grant J et al. AMEE Guide No. 27: Effective educational and clinical supervision. Med Teach 2007; 29: 2–19 doi:10.1080/01421590701210907