Accepted Manuscript online: 2020-08-28 Article published online: 2020-10-06

Diagnostic accuracy of capsule endoscopy compared with colonoscopy for polyp detection: systematic review and meta-analyses

Authors
Tue Kjølhede[1], Anne Mette Ølholm[1], Lasse Kaalby[2,3], Kristian Kidholm[1], Niels Qvist[2], Gunnar Baatrup[2,3]

Institutions
1 Centre for Innovative Medical Technology, Odense University Hospital, Odense, Denmark
2 Department of Surgery, Odense University Hospital, Odense, Denmark
3 Department of Clinical Research, University of Southern Denmark, Odense, Denmark

submitted 12.12.2019
accepted after revision 28.8.2020
published online 28.8.2020

Bibliography
Endoscopy 2021; 53: 713–721
DOI 10.1055/a-1249-3938
ISSN 0013-726X
© 2020. Thieme. All rights reserved.
Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Corresponding author
Gunnar Baatrup, MD, Department of Surgery, Odense University Hospital, House of Research, Baagøes Allé 15, 5700 Svendborg, Denmark
Fax: +45-2877-7610
Gunnar.Baatrup@rsyd.dk

Supplementary material
Supplementary material is available under https://doi.org/10.1055/a-1249-3938

Scan this QR-Code for the author commentary.

ABSTRACT

Background Colon capsule endoscopy (CCE) is a technology that might contribute to colorectal cancer (CRC) screening programs as a filter test between fecal immunochemical testing and standard colonoscopy. The aim was to systematically review the literature for studies investigating the diagnostic yield of second-generation CCE compared with standard colonoscopy.

Methods A systematic literature search was performed in PubMed, Embase, and Web of Science. Study characteristics including quality of bowel preparation and completeness of CCE transits were extracted. Per-patient sensitivity and specificity were extracted for polyps (any size, ≥ 10 mm, ≥ 6 mm) and lesion characteristics. Meta-analyses of diagnostic yield were performed.

Results The literature search revealed 1077 unique papers and 12 studies were included. Studies involved a total of 2199 patients, of whom 1898 were included in analyses. The rate of patients with adequate bowel preparation varied from 40 % to 100 %. The rates of complete CCE transit varied from 57 % to 100 %. Our meta-analyses demonstrated that mean (95 % confidence interval) sensitivity, specificity, and diagnostic odds ratio were: 0.85 (0.73–0.92), 0.85 (0.70–0.93), and 30.5 (16.2–57.2), respectively, for polyps of any size; 0.87 (0.82–0.90), 0.95 (0.92–0.97), and 136.0 (70.6–262.1), respectively, for polyps ≥ 10 mm; and 0.87 (0.83–0.90), 0.88 (0.75–0.95), and 51.1 (19.8–131.8), respectively, for polyps ≥ 6 mm. No serious adverse events were reported for CCE.

Introduction

The American Society for Gastrointestinal Endoscopy [1] and the Danish Health Authority [2] have issued guidelines for screening programs for colorectal cancer (CRC). Several European countries have initiated CRC screening programs, including Denmark, where a national CRC screening program was launched in 2014. The Danish CRC screening program invites individuals aged 50–74 years to submit a fecal immunochemical test (FIT), which is followed by an invitation to standard colonoscopy if the sample contains sufficient traces of occult blood (100 ng Hb/mL buffer). The higher sensitivity of the FIT test compared with former fecal occult blood tests has resulted in a high rate of false-positive test results. In addition to the

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

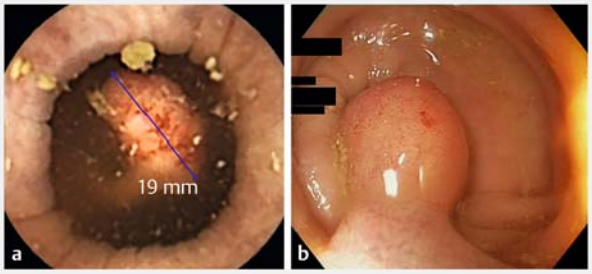
high number of clean colon investigations in FIT-positive individuals, a large proportion of individuals have diminutive (<6 mm) or small (6–9 mm) polyps [3].

The participation rate in the screening program has, during the initial 4 years of screening in Denmark, decreased from 65 % to 61 %, which impacts the efficiency of the program negatively. In addition, the fact that 10 % of FIT-positive individuals refuse further diagnostic tests indicates that colonoscopy is a major cause of nonparticipation. We have previously found indications that the expected and experienced discomfort from colonoscopy is significant, with more than two-thirds of patients indicating medium or severe discomfort/pain, which is reduced by 80 % in camera capsule investigations [4].

Data from the first years of the Danish CRC screening program showed that 7 % of all screened patients had a positive FIT result. Of these, 90 % accepted a standard colonoscopy within 2 months. The results of the colonoscopy showed that 33.7 % had no abnormalities, 5.9 % had cancer, 50.6 % had low risk (36.8 %), medium risk (35.0 %) or high risk (28.2 %) adenomas, while the remaining 9.8 % were unclear and required follow-up investigations such as repeat colonoscopy or computed tomography (CT)-colonography. The frequency of severe complications associated with standard colonoscopy was, according to the Danish Colorectal Cancer Screening Database, 0.2 % [3] and most were related to therapeutic colonoscopies. A later validation review of patient files revealed an incidence of 0.51 % severe complications [5]. The reported frequency of patient-reported complications is up to 23 % within 30 days after colonoscopy, but this also includes minor complications [6].

In order to increase participation and minimize the risk of colonoscopy-related complications in the Danish CRC screening program, alternatives should be considered. One option is to introduce a filter test between FIT and colonoscopy, such as colon capsule endoscopy (CCE) [7] (▶ **Fig. 1**), which, in a screening context, could identify participants with no need for further procedures. The efficacy of CCE as a filter test depends on the diagnostic test accuracy (sensitivity and specificity) of the procedure.

Thus, the aim of this systematic literature review was to describe and summarize the studies investigating the diagnostic test accuracy of second-generation CCE compared with standard colonoscopy in detecting colorectal neoplasia.



▶ **Fig. 1** Images of the same polyp. **a** Colon capsule endoscopy. **b** Standard colonoscopy.

Methods

This review and meta-analysis was performed and reported in accordance with the preferred reporting items for systematic reviews and meta-analyses [8].

Literature search and sorting

A systematic literature search of three scientific databases (PubMed, Embase, and Cochrane Library) was performed on 5 May 2019, and updated on 16 March 2020, to identify research papers investigating the diagnostic test accuracy of CCE compared with colonoscopy. The detailed search strategy is presented in **Appendix 1s** and **Appendix 2s** (see the online-only supplementary material). Briefly, variables were combined technology (second-generation CCE), indication (CRC, neoplasms) and comparison (standard colonoscopy). Titles and abstracts were screened independently by two reviewers in Covidence (a review management tool used for screening and data extraction in literature reviews; Covidence.org, Melbourne, Australia) in order to identify studies directly comparing diagnostic test accuracy of CCE and colonoscopy for patients undergoing both procedures. Full texts were retrieved and read thoroughly by two reviewers. Disagreements were resolved by a third reviewer.

Quality assessment and data extraction

Two reviewers scored all included studies for risk of bias according to the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool [9]. Subsequently, both reviewers extracted study characteristics and data comparing CCE and colonoscopy. Study characteristics included: a) organization, b) participants included, c) age, d) sex, e) indication for colonoscopy, f) regimens used for bowel preparation, g) quality of bowel preparation, and h) percentage of complete CCE transits (CCE egestion within the battery lifespan). The data extracted from the studies were per patient 2 × 2 tables (true positives, false positives, true negatives, and false negatives) and/or sensitivity and specificity for the following outcomes: polyps (any size, ≥ 10 mm, ≥ 6 mm), "patients at risk" (commonly defined as either one polyp ≥ 10 mm or three or more polyps), lesion characteristics (e. g. CRC, laterally spreading tumors).

Additionally, adverse events related to bowel preparation, CCE, and colonoscopy procedures were extracted from the studies, according to each study's definition, and summarized.

Statistics

Statistical analyses were performed in Review Manager 5.3 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark) and Stata 15.1 (StataCorp, College Station, Texas, USA) using the metandi function, which fits a two-level mixed logistic regression model, with independent binomial distributions for the true positives and true negatives conditional on the sensitivity and specificity in each study, and a bivariate model calculated summary points between studies [10, 11]. A random effects model was considered appropriate [12] as a conservative approach considering the difference in inclu-

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

▶ **Table 1** Assessment of risk of bias and applicability concern for all included studies.

	Patient selection		Index test (CCE)		Reference (standard colonoscopy)		Flow and timing
Study	**Bias**	**Applicability**	**Bias**	**Applicability**	**Bias**	**Applicability**	**Bias**
Akyuz 2016 [19]	High	Low	High	High	High	High	Low
Eliakim 2009 [21]	High	Low	Low	Low	High	High	Low
Hagel 2014 [24]	High	Low	Low	Low	High	High	Low
Holleran 2014 [13]	Low	Low	Low	Low	High	High	Low
Igawa 2017 [14]	Low	Low	Low	Low	High	High	Low
Kobaek-Larsen 2018 [15]	Low	Low	Low	Low	High	High	Low
Ota 2017 [20]	High	Low	High	Low	High	High	Low
Parodi 2018 [16]	Low	Low	Low	Low	Low	High	Low
Pecere 2020 [17]	Low	Low	Low	Low	High	High	Low
Rex 2015 [22]	High	Low	Low	Low	High	High	Low
Spada 2011 [23]	High	Low	Low	Low	High	High	Low
Voska 2019 [18]	Low	Low	Low	Low	High	High	Low

CCE, colon capsule endoscopy.

sion criteria for the 12 included studies. Estimates of heterogeneity between studies, τ^2, and Cochran's *Q* test, for diagnostic odds ratios, were calculated in RStudio 1.1.456 (RStudio, Boston, Massachusetts, USA) with R-version 3.5.1 using the mada-package. Summary points and 95 % confidence intervals (CIs) were calculated for sensitivity, specificity, and diagnostic odds ratio for per-patient outcomes that included at least four studies.

Results

Literature search and quality assessment

The literature search and update resulted in 1338 hits, of which 261 were duplicates. After abstract and title screening, 86 full-text papers were retrieved and read thoroughly, and ultimately 12 studies were included in the current review [13–24]. The flowchart is presented in **Fig. 1s**.

The QUADAS-2 assessments of risk of bias and applicability concern are presented in ▶ **Table 1**. Six studies [13–18] had a low risk of bias in relation to patient selection, as either random or consecutive sample of patients were used. Two studies [19, 20] were classified with high risk of bias in relation to the index test (CCE), as descriptions of the blinding process were insufficient. Because of the unproven capability of colonoscopy as a true gold standard, the reference standard (colonoscopy) was classified as high risk of bias in all studies but one [16], which used a segmental unblinding during the colonoscopy procedure.

Study characteristics

Study characteristics are presented in ▶ **Table 2**. Of the 12 studies included, six [16–18, 21–23] were multicenter studies. A total of 2199 patients were included in the 12 studies, of whom 1898 were analyzed in the studies. Indications for inclusion of participants varied, but most were symptomatic patients. Polyethylene glycol was the most frequently used cleansing agent for bowel preparation, whereas several different supplement boosters were used. The rate of adequate bowel preparation varied from 40 % to 100 %. Ten of the 12 studies reported quality of bowel preparation on a 4-point cleansing scale, where "excellent" or "good" were classified as adequate. The four-point scale has previously shown good interobserver agreement [25]. One study reported on a two-point scale (adequate/inadequate) [17], while the final study used a five-point scale [15], with the definition of acceptable being classified as adequate in the current review. The rates of complete CCE transit varied from 57 % to 100 % of all CCEs ingested.

An overview of the extracted per-patient outcomes for each study is presented in **Table 1s**.

Meta-analyses

It was possible to perform meta-analyses on three outcomes: polyps of any size, polyps ≥ 10 mm, and polyps ≥ 6 mm. For polyps of any size, the mean (95 %CI) sensitivity, specificity, and diagnostic odds ratio were 0.85 (0.73–0.92), 0.85 (0.70–0.93), and 30.5 (16.2–57.2), respectively. Corresponding τ^2 was 0.00 and Cochran's *Q* test was nonsignificant ($P = 0.98$). Forest plot and hierarchical summary receiver operating characteristic (HSROC) curve for polyps of any size are presented in ▶ **Fig. 2**.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

▶ **Table 2** Characteristics of all included studies.

Study	Organization	Patients included (analyzed)	Age, mean, years	Female sex, %	Indication(s) for colonoscopy	Bowel preparation	Adequate bowel preparation[1], %	Complete CCE transit[2], %
Akyuz 2016 [19]	Single center	62 (28)	56	66	NA	3 different: PEG vs. NaP vs. PEG + NaP	70	100
Eliakim 2009 [21]	Multicenter	104 (98)	50	34	CRC screening History of polyp/ CRC	PEG + NaP	78	81[3]
Hagel 2014 [24]	Single center	24 (23)	51	42	CRC screening History of polyp/ CRC	PEG + NaP	90	71
Holleran 2014 [13]	Single center	62 (62)	63	45	Positive FIT	PEG + NaP/SPS	92	73
Igawa 2017 [14]	Single center	30 (30)	59	20	Tumor positives	PEG + $NaHCO_3$ + Mg citrate	100	100
Kobaek-Larsen 2018 [15]	Single center	261 (253)	64	42	Positive FIT	PEG + Mg oxide	85	57
Ota 2017 [20]	Single center	20 (20)	71	10	CRC positives	PEG + Mg citrate	40	75
Parodi 2018 [16]	Multicenter	177 (177)	57	55	First-degree relatives to patients with CRC	PEG + NaP	81	100
Pecere 2020 [17]	Multicenter	222 (178)	61	43	Positive FIT	PEG + NaP	88	100
Rex 2015 [22]	Multicenter	884 (695)	57	56	Screening	PEG + Suprep	80	92
Spada 2011 [23]	Multicenter	117 (109)	60	38	CRC screening History of polyp/ CRC	PEG + NaP	81	88[3]
Voska 2019 [18]	Multicenter	236 (225)	59	47	Screening	PEG + NaP	90	89

CCE, colon capsule endoscopy; NA, not available; PEG, polyethylene glycol; NaP, sodium phosphate; CRC, colorectal cancer; FIT, fecal immunochemical test; SPS, sodium picosulfate; Mg, magnesium.
[1] Percentage of patients with bowel preparation rated excellent or good according to the study's applied rating tool.
[2] Percentage of patients with CCE egestion within the battery lifespan.
[3] CCE egested within 8–10 hours, when colonoscopy was performed due to logistical constraints.

For polyps ≥ 10 mm, the mean (95 %CI) sensitivity, specificity, and diagnostic odds ratio were 0.87 (0.82–0.90), 0.95 (0.92–0.97), and 136.0 (70.6–262.1), respectively. Corresponding τ^2 was 0.055 and Cochran's Q was nonsignificant ($P = 0.42$). Forest plot and HSROC curve for polyps ≥ 10 mm are presented in ▶ **Fig. 3**.

For polyps ≥ 6 mm, the mean (95 %CI) sensitivity, specificity, and diagnostic odds ratio were 0.87 (0.83–0.90), 0.88 (0.75–0.95), and 51.1 (19.8–131.8), respectively. Corresponding τ^2 was 1.07 and Cochran's Q was nonsignificant ($P = 0.45$). Forest plot and HSROC curve for polyps ≥ 6 mm are presented in ▶ **Fig. 4**.

Diagnostic yield of other findings

Only the study by Ota et al. [20] reported sensitivity and specificity for detection of cancer. However, as all included patients were CRC positive on inclusion, only sensitivity was reported. The per-patient sensitivity for cancer was 85 % (95 %CI 62 %–97 %).

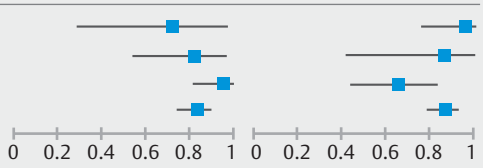
"Patients at risk" was investigated by Kobaek-Larsen et al. [15]. By defining this as at least one polyp ≥ 11 mm or three or more polyps detected, their investigation resulted in a mean sensitivity and specificity of 93 % and 69 %, respectively.

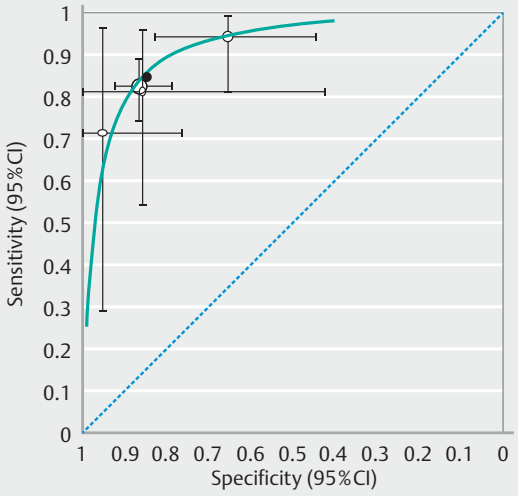
Igawa et al. [14] reported on the capability of CCE to detect laterally spreading tumors compared with colonoscopy. They reported 17 true positives and 4 false positives, thus resulting in a sensitivity of 81 %. They found zero false negatives and nine true negatives corresponding to a specificity of 100 %.

Pecere et al. [17] reported on the capability of CCE to detect advanced neoplasia compared with colonoscopy. Their per-patient results showed mean (95 %CI) sensitivity and specificity for advanced neoplasia of 90.0 (78.8–95.9) and 66.1 (56.7–74.4), respectively, when using a 6 mm cutoff, and 76.7 (63.7–86.2) and 90.7 (83.6–95.0), respectively, when using a 10 mm cutoff.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

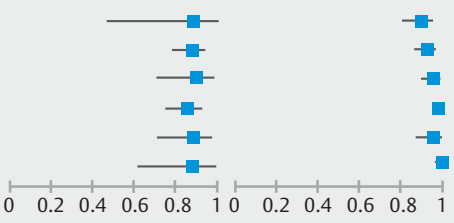
Study	TP	FP	FN	TN	Sensitivity (95 %CI)	Specificity (95 %CI)
Akyuz 2016	5	1	2	20	0.71 [0.29 – 0.96]	0.95 [0.76 – 1.00]
Hagel 2014	13	1	3	6	0,81 [0.54 – 0.96]	0.86 [0.42 – 1.00]
Holleran 2014	34	9	2	17	0.94 [0.81 – 0.99]	0.65 [0.44 – 1.00]
Voska 2019	94	15	20	96	0.82 [0.74 – 0.89]	0.86 [0.79 – 0.92]

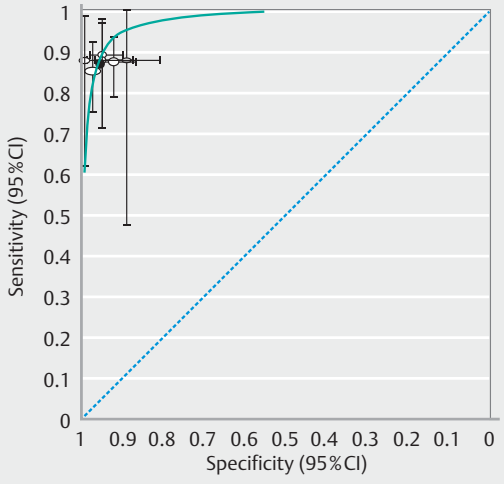




▶ **Fig. 2** Forest plot and hierarchical summary receiver operating characteristic curve for polyps of any size. TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence intervals.

Study	TP	FP	FN	TN	Sensitivity (95 %CI)	Specificity (95 %CI)
Eliakim 2009	7	10	1	1	0.88 [0.47 – 1.00]	0.89 [0.81 – 0.95]
Kobaek-Larsen 2018	81	13	12	147	0,87 [0.79 – 0.93]	0.92 [0.87 – 0.96]
Parodi 2018	24	8	3	142	0.89 [0.71 – 0.98]	0.95 [0.90 – 0.98]
Rex 2015	67	18	12	598	0.85 [0.75 – 0.92]	0.97 [0.95 – 0.98]
Spada 2011	28	4	4	73	0.88 [0.71 – 0.96]	0.95 [0.87 – 0.99]
Voska 2019	14	2	2	207	0.88 [0.62 – 0.98]	0.99 [0.97 – 1.00]

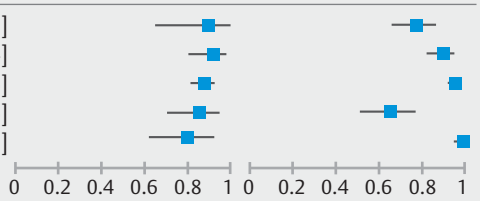


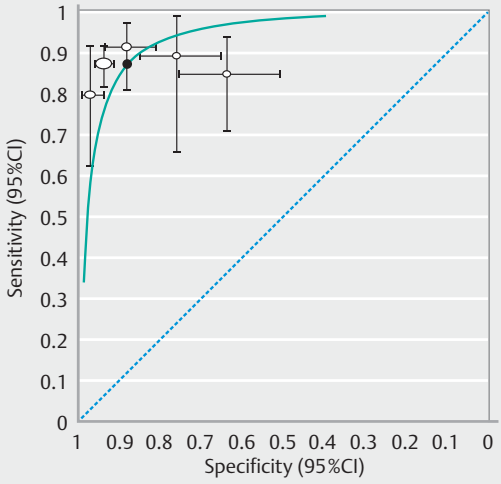


▶ **Fig. 3** Forest plot and hierarchical summary receiver operating characteristic curve for polyps ≥ 10 mm. TP, true positives; FP, false positives; FN, false negative; TN, true negative; CI, confidence intervals.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

Study	TP	FP	FN	TN	Sensitivity (95 %CI)	Specificity (95 %CI)
Eliakim 2009	16	19	2	61	0.89 [0.65, – 0.99]	0.76 [0.65 – 0.85]
Parodi 2018	51	14	5	107	0.91 [0.80 – 0.97]	0.88 [0.81 – 0.94]
Rex 2015	167	30	25	473	0.87 [0.81 – 0.91]	0.94 [0.92 – 0.96]
Spada 2011	38	23	7	41	0.84 [0.71 – 0.94]	0.64 [0.51 – 0.76]
Voska 2019	27	5	7	183	0.79 [0.62 – 0.91]	0.97 [0.94 – 0.99]





▶ **Fig. 4** Forest plot and hierarchical summary receiver operating characteristic curve for polyps ≥ 6 mm. TP, true positives; FP, false positives; FN, false negatives; TN, true negatives; CI, confidence intervals.

Cases with cancer

Eliakim et al. [21] reported one adenomatous polyp with cancer that was identified with both CCE and colonoscopy. Similarly, Holleran et al. [13] identified one patient with cancer with both procedures. Spada et al. [23] observed three cancers (two in the descending colon, one in the sigmoid colon) with both CCE and colonoscopy. Voska et al. [18] observed two cases of carcinoma, which were identified by both procedures.

Kobaek-Larsen et al. [15] identified 11 adenocarcinomas by colonoscopy, 7 of which were also identified by CCE (five as a suspicious large mass, two as polyps). The remaining four cancers were situated distally in the colorectum and were not detected due to the capsule passing the site after battery exhaustion. In all four cases, the CCE investigation was deemed insufficient and led to colonoscopy, during which all cancers were detected. Ota et al. [20] included 20 patients with 21 CRC lesions. Four lesions in three patients were not detected by CCE. All four nondetected lesions were located in the distal (descending) colon, at which point the capsule was not recording. Pecere et al. [17] observed 11 cancers with colonoscopy, but only 10 were detected with CCE, with the undetected cancer being misjudged as a diminutive polyp (≤ 5 mm). Rex et al. [22] observed four cancers with colonoscopy, but only found three by CCE. A review of the capsule recording showed that the cancer was in fact photographed but not detected by the initial review of the recording. In summary, 44 of 54 cancers were detected by CCE, and the cancers missed were located in an area where the CCE was not recording while passing (n = 8), were detected in a second review (n = 1), or the size was misjudged (n = 1).

Parodi et al. [16] reported a mean per-polyp sensitivity of 82.4 % (95 %CI 59.0–93.8) for polyps ≥ 6 mm with high grade dysplasia.

Adverse events

None of the included studies reported any serious adverse events related to the CCE procedure. However, a few studies reported technical problems such as capsules retained in the cecum [18, 21, 23]. The bowel preparation was related to mild adverse events such as vomiting and nausea in several patients [16–18, 20–23, 25], with frequencies reaching 25 % [17]. The colonoscopy procedures were associated with few (14 of 1898 analyzed patients, 0.7 %) moderate to serious adverse events, such as bloating, abdominal pain, rectal bleeding, and bowel perforation, with some of those being related to the polyp removal procedure [15, 16, 18, 22, 23].

Discussion

The main finding of our systematic literature review was that several comparable studies have investigated the sensitivity and specificity of second-generation CCE vs. colonoscopy for various colorectal neoplasia. The meta-analyses showed high per-patient sensitivity and specificity for detecting polyps of any size, polyps ≥ 10 mm, and polyps ≥ 6 mm. We did not identify a sufficient number of studies investigating other outcomes

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

to enable meta-analyses. CCE identified most cases of cancer, but the missed cancers were mostly due to the battery capacity being insufficient to allow recording of the entire gastrointestinal tract.

In terms of moderate and serious adverse events, few (0.7 %) were reported for colonoscopy, but none were reported for CCE suggesting that the capsule is a safer procedure. If CCE was used as a filter test in screening, it would be expected that the complication rate for screening colonoscopy would decrease significantly. Although the most serious complications arise from therapeutic colonoscopies, hundreds of cases of severe complications in the Danish screening program could be avoided annually by introducing CCE. The effect would depend upon the number of colonoscopies that could be avoided by introducing this filter test.

Collectively, the 12 identified studies of second-generation CCE included heterogeneous patients. A meta-analysis comparing first- and second-generation CCE was published by Spada et al. [26] in 2016 and showed progress in sensitivity and specificity by the development of the second generation of capsules; however, it also reported that most previous studies were performed in symptomatic patients. Our systematic review included data from two larger studies that were published after 2016 [15, 16], both of which included asymptomatic patients. The performance of CCE in asymptomatic patients more realistically reflects the performance of CCE in a screening context as a filter test between FIT and colonoscopy.

Overall, the bowel preparation regimens were comparable between the included studies. However, the reported percentage of bowel preparations that were classified as good or excellent varied from 40 % to 100 %. The assessment of bowel preparation is subjective, with interobserver reliability, as assessed by intraclass correlation coefficients, being good but dependent on the grading scale applied [27]. Complete CCE transit similarly varied between 57 % and 100 % in the 12 included studies. The reported per-patient sensitivities and specificities compared the CCE procedures (regardless of complete CCE transit) with the complete colonoscopy procedures, thus reducing the possibility of CCE to detect all polyps and hence skewing the sensitivity measures. However, Kobaek-Larsen et al. [15] performed a subanalysis that included only complete investigations (i. e. patients with complete CCE and colonoscopy), and this showed an improved CCE sensitivity, from 87 % in all patients to 97 % for patients with complete investigations. This highlights the importance of proper bowel preparation to visualize the complete colonic system and to properly compare CCE and colonoscopy. Some studies have investigated the effect of alternative bowel preparation regimens, and a recent study by Ohmiya et al. reported that castor oil in addition to polyethylene glycol increased capsule excretion rates from 81 % to 97 %, and reduced total examination time from 239 minutes to 201 minutes [28].

In this systematic review, the main outcome was per-patient sensitivity and specificity of CCE compared with colonoscopy. Some studies also reported per-polyp detection rates of CCE. This is similarly an important outcome especially in a clinical context, but the exact localization of polyps is difficult to achieve with both colonoscopy and CCE, making comparisons unreliable. Additionally, a per-polyp analysis cannot produce estimates for specificity as there will be zero true negatives.

Another important factor to consider when conducting meta-analyses on diagnostic performance is the quality of the gold standard. A systematic review from 2006 demonstrated a miss rate of 22 % for polyps of any size for repeated colonoscopy [29]; however, a study from 2012, which performed back-to-back colonoscopy, demonstrated miss rates of polyps, adenomas, and advanced adenomas of 16.8 %, 17 %, and 5.4 %, respectively [30]. Accordingly, in the study by Rex et al. [22], 52 polyps were identified by CCE but missed by colonoscopy. Of these, 22 % of the missed polyps were verified by a later repeat colonoscopy. These findings were verified by Kobaek-Larsen et al. [15]. Collectively, this demonstrates that colonoscopy may not be the most appropriate gold standard, and that some of the reported false-positive CCE investigations in the included studies might in fact have been true positives; thus, our results might underestimate the true sensitivity of CCE and overestimate the true specificity of colonoscopy. Additionally, some of the included studies applied strict criteria with regard to size estimates of polyps. For instance, Parodi et al. [16] considered patients as false positives if CCE estimated a lesion ≥ 6 mm but colonoscopy estimated the lesion to be ≤ 5 mm, which again skews the results negatively for CCE. However, studies show that even in situ measurements of polyp size vary between trained observers [31].

The possible implementation of CCE in national screenings programs is dependent on the diagnostic test accuracy of CCE, but also the cost of the capsule and the analysis of the data obtained. Artificial intelligence is improving rapidly in image diagnostics, although the quality of studies is still inadequate [32]. For CCE, a study was published recently by Blanes-Vidal et al. [33] and showed an accuracy of 96.4 %, sensitivity of 97.1 %, and specificity of 93.3 % for an autonomous detection algorithm compared with trained nurses and gastroenterologists in the detection and localization of colorectal polyps. The authors further indicated that polyp size estimation can be improved by artificial intelligence compared with endoscopist evaluation, possibly contributing to more correct allocation of patients to further treatment and follow-up [33]. If the data analysis of the obtained images to a large extent can be performed by artificial intelligence, the complete cost of a capsule procedure could be reduced significantly, and the implementation in screening programs might become more feasible. Algorithms might also very soon enable real-time diagnostics, enabling a subsequent therapeutic colonoscopy to be performed immediately during the same session with no further bowel preparation, further reducing costs.

Acceptability of CCE is important to consider in a national screening context. Ojidu et al. [34] observed significantly better patient experience with CCE compared with colonoscopy and CT-colonography. Surprisingly, they also observed that only 85.7 % of the CCE patients were willing to undergo the same test again, compared with 93.6 % and 96.1 % for colonoscopy and CT-colonography patients, respectively. We investigated this in a previous publication and found support for this

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

statement [4]. In addition to acceptability, logistics is another concern. Ideally, colonoscopy should be performed during the same session with no further bowel cleansing required. This would be logistically challenging as it demands immediate reading of the CCE examination and subsequent access to colonoscopy. The effect of patient compliance needs to be investigated further in a randomized setting.

The examination completion rate requires improvement if CCE is to earn potential utility in a screening context, and studies directly comparing procedures on long-term end points should be performed rather than focusing on more aggressive bowel preparation in an attempt to improve CCE completion rates. An alternative is to prioritize technical improvements, such as improved battery capacity.

The cost of CCE examinations is also important in the screening context. Most studies show that approximately half of the CCE screening patients undergo subsequent colonoscopy. Accordingly, one could argue that the cost of CCE should be half of the cost for a colonoscopy to make expenditures even. However, the cost of screening programs is usually determined as a price per life-year gained. To our knowledge, no study has yet been published on this topic.

Another important question to consider in a national screening context, is how to handle diminutive (< 6 mm) and small (6–9 mm) polyps. These polyps are considered incidental findings as they very rarely bleed and the frequency is not very different in FIT-negative individuals. There is widespread understanding that the resection of diminutive polyps does not add to the positive effects of the screening program, and the effect of removing small polyps is highly questionable. The concepts of "diagnose and leave behind" or "resect and discard" have been described as a potential approach to manage diminutive polyps [35].

Although the second generation of colon capsules has been available for more than a decade, the modality has not gained wide use. The main reasons are logistics in delivering capsules to the patients and reading the obtained records, costs, incomplete examinations, and the number of patients who undergo a subsequent colonoscopy. The development of artificial intelligence for a computer-based evaluation of records, changes in guidelines for the handling of small polyps, lower capsule prices, and longer capsule battery life may be some of the solutions.

In conclusion, second-generation CCE has a high sensitivity (means and 95 %CI ranging from 0.85–0.87 and 0.73–0.90, respectively) and specificity (0.85–0.95 and 0.70–0.97, respectively) for per-patient polyps compared with colonoscopy. However, the relatively high rate of incomplete CCE transit and issues regarding bowel preparation quality indicate that improvements are necessary before clinical implementation of CCE in CRC screening becomes feasible.

Acknowledgments

This study was financially supported by Odense University Hospital. Camera capsules were provided, free of charge, by Medtronic, for research purposes.

Competing interests

The authors declare that they have no conflicts of interest.

References

[1] Davila RE, Rajan E, Baron TH et al. ASGE guideline: colorectal cancer screening and surveillance. Gastrointest Endosc 2006; 63: 546–557

[2] Sundhedsstyrelsen, Anbefalinger vedrørende screening for tyk- & endetarmskræft. Sundhedsstyrelsen. 2012: Available at (Accessed 11 September 2020): https://www.sst.dk/-/media/Udgivelser/2012/Publ2012/Anbefalninger-vedr%C3%B8rende-screening-for-tyk–og-endetarmskr%C3%A6ft.ashx

[3] Årsrapport 2017. Danish CRC-screening Database; 2018: Available at (Accessed 11 September 2020): https://dccg.dk/wp-content/uploads/2019/03/DCCG-%C3%85rsrapport-2017v2.pdf

[4] Thygesen MK, Baatrup G, Petersen C et al. Screening individuals' experiences of colonoscopy and colon capsule endoscopy; a mixed methods study. Acta Oncol 2019; 58: (Suppl. 01): S71–S76

[5] Mikkelsen EM, Thomsen MK, Tybjerg J et al. Colonoscopy-related complications in a nationwide immunochemical fecal occult blood test-based colorectal cancer screening program. Clin Epidemiol 2018; 10: 1649–1655

[6] Steffenssen MW, Al-Najami I, Baatrup G. Patient-reported minor adverse events after colonoscopy: a systematic review. Acta Oncol 2019; 58: (Suppl. 01): S22–S28

[7] 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

[8] Moher D, Liberati A, Tetzlaff J et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009; 6: e1000097

[9] Whiting PF, Rutjes AW, Westwood ME et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; 155: 529–536

[10] Reitsma JB, Glas AS, Rutjes AW et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005; 58: 982–990

[11] Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. J Clin Epidemiol 2006; 59: 1331–1332

[12] Lee J, Kim KW, Choi SH et al. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers – Part II. Statistical methods of meta-analysis. Korean J Radiol 2015; 16: 1188–1196

[13] Holleran G, Leen R, O'Morain C et al. Colon capsule endoscopy as possible filter test for colonoscopy selection in a screening population with positive fecal immunology. Endoscopy 2014; 46: 473–478

[14] Igawa A, Oka S, Tanaka S et al. Evaluation for the clinical efficacy of colon capsule endoscopy in the detection of laterally spreading tumors. Digestion 2017; 95: 43–48

[15] Kobaek-Larsen M, Kroijer R, Dyrvig AK et al. Back-to-back colon capsule endoscopy and optical colonoscopy in colorectal cancer screening individuals. Colorectal Dis 2018; 20: 479–485

[16] Parodi A, Vanbiervliet G, Hassan C et al. Colon capsule endoscopy to screen for colorectal neoplasia in those with family histories of colorectal cancer. Gastrointest Endosc 2018; 87: 688

[17] Pecere S, Senore C, Hassan C et al. Accuracy of colon capsule endoscopy for advanced neoplasia. Gastrointest Endosc 2020; 91: 406

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

[18] Voska M, Zavoral M, Grega T et al. Accuracy of colon capsule endoscopy for colorectal neoplasia detection in individuals referred for a screening colonoscopy. Gastroenterol Res Pract 2019; 2019: 5975438

[19] Akyuz U, Yilmaz Y, Ince AT et al. Diagnostic role of colon capsule endoscopy in patients with optimal colon cleaning. Gastroenterol Res Pract 2016; 2016: 2738208

[20] Ota Y, Yamada A, Kobayashi Y et al. Diagnostic capability of colon capsule endoscopy for advanced colorectal cancer: a pilot study. Dig Endosc 2017; 29: 695–701

[21] 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

[22] Rex DK, Adler SN, Aisenberg J et al. Accuracy of capsule colonoscopy in detecting colorectal polyps in a screening population. Gastroenterology 2015; 148: 948–967

[23] Spada C, Hassan C, Munoz-Navas M et al. Second-generation colon capsule endoscopy compared with colonoscopy. Gastrointest Endosc 2011; 74: 581–589

[24] Hagel AF, Gabele E, Raithel M et al. Colon capsule endoscopy: detection of colonic polyps compared with conventional colonoscopy and visualization of extracolonic pathologies. Can J Gastroenterol Hepatol 2014; 28: 77–82

[25] Leighton JA, Rex DK. A grading scale to evaluate colon cleansing for the PillCam COLON capsule: a reliability study. Endoscopy 2011; 43: 123–127

[26] Spada C, Pasha SF, Gross SA et al. Accuracy of first- and second-generation colon capsules in endoscopic detection of colorectal polyps: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2016; 14: 1533–1543

[27] Kastenberg D, Bertiger G, Brogadir S. Bowel preparation quality scales for colonoscopy. World J Gastroenterol 2018; 24: 2833–2843

[28] Ohmiya N, Hotta N, Mitsufuji S et al. Multicenter feasibility study of bowel preparation with castor oil for colon capsule endoscopy. Dig Endosc 2019; 31: 164–172

[29] van Rijn JC, Reitsma JB, Stoker J et al. Polyp miss rate determined by tandem colonoscopy: a systematic review. Am J Gastroenterol 2006; 101: 343–350

[30] Ahn SB, Han DS, Bae JH et al. The miss rate for colorectal adenoma determined by quality-adjusted, back-to-back colonoscopies. Gut Liver 2012; 6: 64–70

[31] Buijs MM, Steele RJC, Buch N et al. Reproducibility and accuracy of visual estimation of polyp size in large colorectal polyps. Acta Oncol 2019; 58: (Suppl. 01): S37–S41

[32] Liu X, Faes L, Kale AU et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health 2019; 1: e271–e297

[33] Blanes-Vidal V, Baatrup G, Nadimi ES. Addressing priority challenges in the detection and assessment of colorectal polyps from capsule endoscopy and colonoscopy in colorectal cancer screening using machine learning. Acta Oncol 2019; 58: (Suppl. 01): S29–S36

[34] Ojidu H, Palmer H, Lewandowski J et al. Patient tolerance and acceptance of different colonic imaging modalities: an observational cohort study. Eur J Gastroenterol Hepatol 2018; 30: 520–525

[35] Ferlitsch M, Moss A, Hassan C et al. Colorectal polypectomy and endoscopic mucosal resection (EMR): European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline. Endoscopy 2017; 49: 270–297

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.