Educational interventions are associated with improvements in colonoscopy quality indicators: a systematic review and meta-analysis

Authors

Natalia S. Causada-Calo*[,1], Emmanuel I. Gonzalez-Moreno*[,2,3], Kirles Bishay[1], Risa Shorr[4], Catherine Dube[5,6], Steven J. Heitman[2,3], Robert J. Hilsden[2,3], Alaa Rostom[5,6], Catharine Walsh[7,8,9], John T. Anderson[10], Rajesh N. Keswani[11], Michael A. Scaffidi[1], Samir C. Grover[1,12], Nauzer Forbes[2,3]

Institutions

1 Division of Gastroenterology, St. Michael's Hospital, Toronto, Canada
2 Division of Gastroenterology and Hepatology, University of Calgary, Calgary, Canada
3 Department of Community Health Sciences, University of Calgary, Calgary, Canada
4 Learning Services, The Ottawa Hospital, Ottawa, Ontario, Canada
5 Division of Gastroenterology, Department of Medicine, University of Ottawa, Ottawa, Canada
6 Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Canada
7 Division of Gastroenterology, Hepatology, and Nutrition, Learning Institute, and Research Institute, Hospital for Sick Children, Toronto, Canada
8 The Wilson Centre, University of Toronto, Toronto, Canada
9 Department of Pediatrics, University of Toronto, Toronto, Canada
10 Department of Gastroenterology, Gloucestershire Hospitals NHSFT, Gloucester, UK
11 Division of Gastroenterology, Feinberg School of Medicine, Northwestern University, Chicago, USA
12 Li Ka Shing Knowledge Institute, University of Toronto, Toronto, Ontario, Canada

submitted 8.4.2020
accepted after revision 25.6.2020

Bibliography
Endoscopy International Open 2020; 08: E1321–E1331
DOI 10.1055/a-1221-4922
ISSN 2364-3722

Corresponding author
Nauzer Forbes, MD, MSc, Clinical Assistant Professor of Medicine, Division of Gastroenterology, University of Calgary, TRW 6D19, 3280 Hospital Drive NW, Calgary, AB.T2N 4Z6, Canada
Fax: +403.592.5090
nauzer.forbes@ucalgary.ca

Supplementary material is available under https://doi.org/10.1055/a-1221-4922

ABSTRACT

Background and study aims The quality of screening-related colonoscopy depends on several physician- and patient-related factors. Adenoma detection rate (ADR) varies considerably between endoscopists. Educational interventions aim to improve endoscopists' ADRs, but their overall impact is uncertain. We aimed to assess whether there is an association between educational interventions and colonoscopy quality indicators.

Methods A comprehensive search was performed through August 2019 for studies reporting any associations between educational interventions and any colonoscopy quality indicators. Our primary outcome of interest was ADR. Two authors assessed eligibility criteria and extracted data independently. Risk of bias was also assessed for included studies. Pooled rate ratios (RR) with 95 % confidence intervals (CI) were reported using DerSimonian and Laird random effects models.

Results From 2,253 initial studies, eight were included in the meta-analysis for ADR, representing 86,008 colonoscopies. Educational interventions were associated with improvements in overall ADR (RR 1.29, 95 % CI 1.25 to 1.42, 95 % prediction interval 1.09 to 1.53) and proximal ADR (RR 1.39, 95 % CI 1.29 to 1.48), with borderline increases in withdrawal time, ([WT], mean difference 0.29 minutes, 95 % CI – 0.12 to 0.70 minutes). Educational interventions did not affect cecal intubation rate ([CIR], RR 1.01, 95 % CI 1.00 to

* These authors contributed equally.

1.01). Heterogeneity was considerable across many of the analyses.

Conclusions Educational interventions are associated with significant improvements in ADR, in particular, proximal ADR, and are not associated with improvements in WT or CIR. Educational interventions should be considered an important option in quality improvement programs aiming to optimize the performance of screening-related colonoscopy.

Introduction

In 2020, it is estimated that nearly 150,000 individuals will be newly diagnosed with colorectal cancer (CRC) in the United States alone, with over 50,000 expected CRC-related deaths [1]. CRC is a disease process that is appropriate for employing population-based screening, given that its natural history typically involves a slow progression from adenoma to cancer [2]. Colonoscopy has been shown to reduce CRC incidence and mortality in a cost-effective fashion [3, 4], given its capacity to both identify and remove adenomatous polyps. It can thus potentially act either as a primary screening modality or a primary method of following up on other abnormal screening tests.

To assess and optimize the overall quality of screening-related colonoscopy, several surrogate indicators have been widely adopted, including cecal intubation rate (CIR) and adenoma detection rate (ADR) [5, 6]. There is a well-established relationship between higher ADR and lower incidence of post-colonoscopy CRC (PCCRC) [7]. Rates of PCCRC can vary depending on the methodology used for their calculation, but are thought to range from 3 % to 13 % with an estimated average of 7.4 % [8, 9]. Several screening programs mandate minimum ADR benchmarks of ≥ 25 % for screening colonoscopy [5, 10], with some advocating for higher targets depending on the population(s) being screened [11].

Despite efforts to improve colonoscopy quality, wide variations in endoscopists' ADRs exist [12 – 14]. Suboptimal technique and inadequate withdrawal times (WTs) are considered to be major factors responsible for this disparity [15]. Several interventions have been studied that aim to improve ADR, ranging from medical devices [16] to optimized endoscopy curricula for trainees [17]. Of particular interest are educational interventions specifically targeted at independently practicing endoscopists. Despite planned systematic performance improvement interventions [18], early evidence failed to demonstrate significant improvements in ADR, leading some to conclude that educational interventions do not improve colonoscopy quality. However, more recent evidence reported that the implementation of a bundle of educational interventions was associated with significant ADR improvements in the poorest performers [19].

We performed a systematic review and meta-analysis to determine whether there is an association between educational interventions and improvements in ADR or any other colonoscopy quality indicators.

Methods

Overview

Our systematic review was conducted and reported according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement recommendations [20] and Meta-analysis of Observational Studies in Epidemiology (MOOSE) statement recommendations. A detailed PRISMA checklist is provided in **Supplementary Table 1**. The protocol for this review was registered on PROSPERO (CRD42019149683). Our primary objective was to determine if there is an association between educational interventions for endoscopists and improvement in ADR. Our secondary objectives were to evaluate whether endoscopist educational interventions are associated with improvements in other colonoscopy quality indicators: polyp detection rate (PDR), advanced neoplasm detection rate (ANDR), proximal ADR (pADR), CIR and WT.

Search strategy

A comprehensive search strategy was developed by members of the study team in conjunction with a health research librarian. We searched MEDLINE, EMBASE (Excerpta Medica Database), Google Scholar, and CENTRAL (Cochrane Central Registry of Controlled Trials) from inception of the databases through August 31, 2019. Our full search strategy is provided along with additional gray literature searches performed in **Appendix 1**.

Eligibility criteria

A study was eligible for inclusion if it was a cohort study, quasi-experimental study or clinical trial, it was published in English as either an abstract or manuscript, it assessed the effect of an educational intervention targeting colonoscopy (including live lectures, slide decks, video tutorials, online training modules, individualized assessment and optimization, and skills enhancement and training courses), *and* it reported on at least one colonoscopy quality indictor (including ADR, PDR, ANDR, pADR, CIR, or WT).

A study was excluded if it reported on data that overlapped with another published study, in part or in whole (in these cases, the study with longer follow up or more complete data was included), it assessed the effect of educational interventions on the performance of trainees, or it assessed only the effect of audit and feedback or other interventions without a targeted educational intervention.

Study selection

Following removal of duplicates, citations were imported into Rayyan (M Ouzzani, Qatar Computing Research Institute, HBKU, Doha, Qatar). All abstracts were screened independently by two reviewers (EGM, KB). In the case of disagreements, a third author (NF) reviewed the study and consensus was achieved. The full-length texts of selected abstracts were retrieved and reviewed.

Data extraction and study quality

A data abstraction form was designed a priori to collect data from each included study. Two reviewers (EGM, KB) independently extracted pre-established data points, in addition to performing assessments of bias and overall study quality. The risk of bias in individual studies was determined using the Newcastle-Ottawa Scale (NOS) for non-randomized studies [21]. Interreviewer discrepancies in data abstraction were resolved by consensus after input of a third author (NF). When studies met inclusion criteria but had insufficient data to be included in the quantitative meta-analysis, we attempted to contact corresponding authors to obtain missing information; if unsuccessful, the study was summarized qualitatively only. We used the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) system to assess the certainty of the evidence according to study design, consistency, directness, imprecision and reporting bias [22].

Outcomes

The primary outcome of our study was ADR. Secondary outcomes of interest were PDR, ANDR, pADR, WT and CIR.

Statistical analysis

Rate ratios (RR) with their respective 95 % confidence intervals (CI) were pooled and presented in Forest plots to estimate the effect of educational interventions on outcomes. Whenever a randomized controlled trial (RCT) also presented preintervention and post-intervention data, we: a) pooled these data along with the meta-analysis of other quasi-experimental observational studies; and b) analyzed the randomized data separately. We defined RR as the ratio of the post-intervention quality indicator divided by the pre-intervention value. For continuous variables such as WT, we calculated mean differences in preintervention and post-intervention measures. We used DerSimonian and Laird random effects models to account for expected heterogeneity across study designs. In addition, the 95 % prediction interval for the primary outcome of interest was calculated [23]. χ^2 tests and I^2 statistics were calculated as a measure of between study heterogeneity. I^2 values of 0 – 30 % were regarded as possibly unimportant, values of 30 % to 50 % were regarded to represent moderate heterogeneity, values of 50 % to 75 % were regarded to represent substantial heterogeneity, and values > 75 % were regarded to represent considerable heterogeneity [24]. Funnel plots as well as Egger's and Begg's tests were used to assess publication bias [25, 26].

To assess other potential sources of heterogeneity, we performed several subgroup analyses, including by study design (RCTs versus observational studies), number of centers (single-center versus multicenter), and publication type (conference abstract versus published manuscript). Subgroup analyses were also performed by indication for colonoscopy (primarily screening-related versus mixed populations), education type (hands-on versus didactic training programs), Endoscopic Quality Improvement Program (EQUIP)-based [27] versus other strategies, and presence of lag time (any versus none). Lag time was defined as any time period between the intervention and outcome measurement. Meta-regression analyses were not performed given that there were fewer than 10 studies in the analysis for the primary outcome [24].

We conducted sensitivity analyses whereby each study was removed in turn and whereby fixed effects models were used rather than random effects models. Statistical analyses were performed using STATA version 14.2 (StataCorp, College Station, Texas, United States) and Revman 5.3 (Cochrane Collaboration).

Results

Study selection

A PRISMA flowchart summarizing the overall search results and study selection process is presented in **Supplementary Fig. 1**. A total of 2,253 citations were identified from the search strategy, without any additional studies identified through manual searches. Of these, 30 full-text articles were reviewed. Eight studies were included in the meta-analysis for the primary outcome. An additional two studies did not contain sufficient data to be quantitatively analyzed despite attempts to contact study authors; thus, these were summarized qualitatively.

Study characteristics and quality

Baseline characteristics of the studies included in the meta-analysis are summarized in ▶**Table 1**. Seven studies were performed in North America and one was performed in Europe. Included studies were published between 2010 and 2019. Three were RCTs [27 – 29], four employed quasi-experimental designs (pre-comparisons and post-comparisons) [30 – 33] and one was a retrospective cohort with pre-comparisons and post-comparisons [34]. A summary of interventions and outcomes from studies included in the meta-analysis is provided in ▶**Table 2**.

Summaries of quality assessments are provided in ▶**Table 1**, with full assessments provided in **Supplementary Table 2** and **Supplementary Fig. 2**. Study quality was high for fully published studies, with a mean NOS of 8.25. Assessments of the certainty of the evidence according to the GRADE approach [22] are summarized in ▶**Table 3**. Summaries of studies included in the systematic review, but not quantitatively analyzed via meta-analysis, are provided in **Supplementary Table 3**.

Adenoma detection rate

Meta-analysis of eight studies compared the ADR pre-education and post-education as RRs and included 86,008 colonoscopies. The pooled baseline ADR was 26.5 % and the post-intervention ADR was 35.4 %. Educational interventions were associated with a 29 % relative increase in ADR (RR 1.29, 95 % CI

▶ **Table 1** Summary of characteristics of studies included in the meta-analysis.

Author, year	Study type	Country	Number of study sites	Endoscopists (N =), specialty, practice type	Colonoscopies (N =) pre-/post-intervention	Patient sex (% male)	Median patient age	Indication (% screening-related)	Study quality
Berger 2017 [33]	OBS	USA	1	11, 100 % GI, academic	1,113/849	N/R	N/R	N/R	N/A*
Coe 2013 [27]	RCT	USA	1	15, 100 % GI, academic	602/520	51	63	42	NOS-8
Corley 2019 [32]	OBS	USA	20	85, specialty N/R, setting N/R	12,266/20,897	49	63	22	N/A*
Evans 2019 [34]	OBS	Canada	1	14, specialty N/R, academic	833/850				
Hall 2010 [31]	OBS	USA	1	11, 100 % GI, academic	550/413	48	53	100	N/A*
Kaminski 2016 [29]	RCT	Poland	38	38, specialty N/R, setting: National CRC Screening Program	14,264/10,615	39	57	100	NOS-9
Keswani 2015 [30]	OBS	USA	1	20, specialty N/R, setting N/R	2,444/3,639	N/R	N/R	100	NOS-8
Wallace 2017 [28]	RCT	USA	9	N/R	7,480/8,673	59	47	70	NOS-8

OBS, observational study; RCT, randomized controlled trial; GI, gastroenterologist; Sx, surgeon; NOS, Newcastle-Ottawa Scale [21]; N/R = not reported
* conference abstract.

1.22 to 1.37, 95 % prediction interval 1.09 to 1.53) as shown in ▶ **Fig. 1**. There was considerable heterogeneity between the eight included studies, demonstrated by an I^2 value of 82.96 % (▶ **Fig. 1**).

Two additional studies met inclusion criteria for the primary outcome but did not contain sufficient data to be included in the meta-analysis, even after attempts to contact corresponding author(s) were made. Both studies reported significant improvements in ADR following multi-level educational interventions [35, 36].

Polyp detection rate

Three studies representing 19,237 colonoscopies compared PDR pre- and post-educational interventions. At baseline, PDR was 49.4 %, whereas after the intervention it increased to 61.1 %. We found that educational interventions were associated with a 23 % relative increase in PDR (RR 1.23, 95 % CI 1.19 to 1.27), as demonstrated in ▶ **Fig. 2a**. There was low heterogeneity between the included studies (I^2 value of 11.21 %).

Proximal adenoma detection rate

Two studies reported pADR before and after educational interventions, representing 26,001 colonoscopies. Prior to the intervention, the pADR was 9.3 %, whereas after the intervention it increased to 13.10 %. Educational interventions were associated with a 39 % relative increase in pADR (RR 1.39, 95 % CI 1.29 to 1.48) as displayed in ▶ **Fig. 2b**. There was no heterogeneity between the studies (I^2 value of 0 %).

Withdrawal time

Four studies representing 50,106 colonoscopies that compared WT before and after educational interventions were included in the meta-analysis. WT was relatively homogeneously defined across studies, being either calculated using negative procedures only, or using all procedures with a timer used to remove period(s) spent on any intervention(s) performed. There were no significant differences in WT (MD 0.29 minutes, 95 % CI – 0.12 to 0.70 minutes) as shown in ▶ **Fig. 2c**. There was considerable heterogeneity between these studies (I^2 value of 93.96 %). Of note, the WTs in the majority of the included studies exceeded those proposed by guidelines [5, 6], ranging from 8.6 to 12.1 minutes in the preintervention arm and from 8.4 to 12.5 minutes in the post-intervention arm.

Cecal intubation rate

Two studies reported on CIR preintervention and post-intervention, representing 25,568 colonoscopies. There were no significant changes in CIR before and after educational interventions, as shown in ▶ **Fig. 2d** (RR 1.00, 95 % CI 1.00 to 1.01). There was no heterogeneity between the included studies (I^2 value of 0.08 %). Of note, the pooled CIR met recommended targets [5, 6] in both the pre- and post-intervention periods, at 96.2 % and 96.4 %, respectively.

▶ **Table 2** Summary of interventions and outcomes from studies included in the meta-analysis.

Author, year	Description of educational intervention	Number of sessions	Lag time after intervention*	Post-intervention observation period	Preintervention ADR	Post-intervention ADR	Other outcomes reported
Berger 2017 [33]	One-hour slide show presentation/ lecture -focusing on improving ADR (EQUIP I/II intervention).	One	None	N/R	33.0	41.9	PDR, SPDR
Coe 2013 [27]	EQUIP I: two slide show presentations that included videos, images, and reference material along with pre- and post- tests (each session duration of approximately 1 hour). *First session*: methods and technical aspects; lesion recognition (particularly flat lesions). *Second session*: pre- and post-test on neoplastic vs. non-neoplastic lesions and advanced imaging modalities.	Two	4 months	7 months	36.0	47.0	PDR, ANDR, pADR
Corley 2019 [32]	30-minute interactive online training module on polyp identification, cleaning/washing techniques, and colonoscopy quality, combined with feedback on ADR.	One	None	24 months	31.5	37.4	None
Evans 2019 [34]	Colonoscopy Skills Improvement (CSI) program, consisting of one day live endoscopy sessions, with two certified faculty teaching up to three 3 endoscopists per session.	One	None	8 months	31.8	35.3	WT, CIR
Hall 2010 [31]	Departmental education regarding current national recommendations regarding withdrawal times and expected detection rates.	One	28 months	3 months	22.0	34.0	WT
Kaminski 2016 [29]	Train-Colonoscopy-Leaders (TCL) cours, comprising three phases. *Phase I*: a 2-hour assessment visit by endoscopy nurses (10 colonoscopies) and two-day training by UK trainers (skills improvement, training the trainer, leadership training). Phase II: 2-day hands on training. Phase III: repeat previous nurse assessments (10 colonoscopies); evaluation of first 30 colonoscopies with feedback.	Two	None	18 months	18.4	24.1	PDR, SPDR, CIR, pADR
Keswani 2015 [30]	Physician report cards containing endoscopists' and institutional ADRs and WT, as well as the institutional mean ADR and ADRs of the 10th and 90th percentile. Concurrent educational meeting detailing the data supporting ADR and its relationship with interval CRC cancer and the report card methodology.	One	12 months	6 months	28.0	39.0	WT
Wallace 2017 [28]	One-hour lecture focusing on improving adenoma detection (EQUIP I/II intervention) followed by: 1–2-h review session, identification of low performers, discussion of obstacles to high quality colonoscopy. In addition, optional one-on-one proctoring is offered as well as telephone calls to discuss progress. EQUIP posters are posted in endoscopy units.	One	N/R	N/R	31.0	42.0	PDR, ANDR, CDR, WT

ADR, adenoma detection rate; CIR, cecal intubation rate; WT, withdrawal time; SPDR, sessile polyp detection rate; PDR, polyp detection rate; ANDR, advanced neoplasia detection rate; pADR, proximal adenoma detection rate; N/R, not reported.
* Lag time refers to time between intervention and start of post-intervention measurement of outcome(s).

▶ **Table 3** GRADE summary of effects of feedback interventions on colonoscopy quality indicators [22].

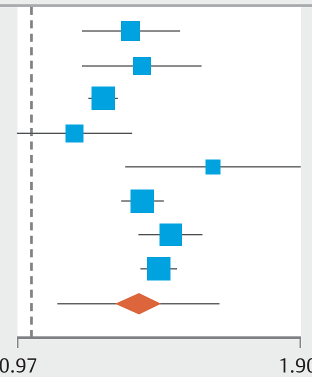
Outcomes	Anticipated absolute effects[1] (95 % CI)		Relative effect (95 % CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with control	Risk with Educational interventions				
Adenoma detection rate – all studies as observational (ADR – Obs)	265 adenomas detected per 1,000 colonoscopies	341 per 1,000 (323 to 362)	Rate ratio 1.29 (1.22 to 1.37)	86008 (8 observational studies)	⊕⊕○○ LOW	Educational interventions likely result in an increase in adenoma detection rate – all studies as observational.
Adenoma detection rate – only non-randomized studies (ADR)	308 adenomas detected per 1,000 colonoscopies	391 per 1,000 (357 to 431)	Rate ratio 1.27 (1.16 to 1.40)	43854 (5 observational studies)	⊕○○○ VERY LOW[2]	Educational interventions may result in an increase in adenoma detection rate.
Adenoma detection rate – only RCTs (ADR-RCTs)	267 adenomas detected per 1,000	315 per 1,000 (283 to 350)	Rate ratio 1.18 (1.06 to 1.31)	25791 (3 RCTs)	⊕⊕⊕○ MODERATE	Educational interventions likely increase adenoma detection rate – only RCTs.
Polyp Detection Rate (PDR)	494 polyps detected per 1,000 colonoscopies	608 per 1,000 (593 to 628)	Rate ratio 1.23 (1.20 to 1.27)	19237 (3 observational studies)	⊕○○○ VERY LOW[2]	Educational interventions may increase polyp detection rate but the evidence is very uncertain.
Proximal adenoma detection rate (pADR)	93 proximal adenomas detected per 1,000 colonoscopies	130 per 1,000 (120 to 138)	Rate ratio 1.39 (1.29 to 1.48)	26001 (2 observational studies)	⊕○○○ VERY LOW[2]	Educational interventions may increase proximal adenoma detection rate but the evidence is very uncertain.
Withdrawal time (WT)	The mean withdrawal time was 10.5 minutes	MD 0.29 minutes higher (0.18 lower to 0.76 higher)	–	48393 (4 observational studies)	⊕○○○ VERY LOW	Educational interventions may increase withdrawal time but the evidence is very uncertain.
Cecal intubation rate (CIR)	962 per 1,000 colonoscopies	962 per 1,000 (962 to 971)	Rate ratio 1.00 (1.00 to 1.01)	26562 (2 observational studies)	⊕○○○ VERY LOW	Educational interventions may have little to no effect on cecal intubation rate but the evidence is very uncertain.

ADR, adenoma detection rate; RR, rate ratio; CI, confidence interval; MD, mean difference; OBS, observational studies
GRADE Working Group grades of evidence:
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect
[1] The risk in the intervention group (and its 95 % confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95 % CI).
[2] a. All the observational studies are quasi-experimental with before and after comparisons. b. The heterogeneity, both statistical and clinical, is substantial and cannot be fully explained. c. The overlap between the confidence intervals is very limited. d. The confidence in the estimate is low (the 95 % CI for the pooled estimate is wide and/or crosses the line of no effect).

Subgroup analyses

Lag times (times between the intervention and the start of outcome measurement) were clearly reported in five studies and varied between 6 and 28 months. The improvement in ADR was slightly less pronounced in studies reporting any lag time following educational interventions compared with studies with no lag time or no lag time reported (RR 1.28, 95 % CI 1.19 to 1.38 and RR 1.34, 95 % CI 1.28 to 1.41, respectively). However, this difference was not statistically significant ($p = 0.32$). The first subgroup, however, had substantial inter-study heterogeneity (I^2 82 %) whereas the second had low heterogeneity (I^2 11 %). There were no significant differences in ADR improvements between studies in screening and mixed populations, hands-on versus didactic training programs, or EQUIP-based versus other strategies. Interestingly, heterogeneity was absent between the EQUIP-based studies (I^2 0 %) and when randomized trials were grouped together using before-and-after data (I^2 0 %). There were also no significant differences in subgroup analyses by study design (RCT versus observational), number of centers (single-center versus multicenter), or publication type (published manuscript versus conference abstract). The inter-study heterogeneity was slightly reduced within the subgroup containing only single-center studies (I^2 value of 64 %), and in the subgroup containing only full text publications (I^2 59 %). When

Study	Educational Int. Adenoma (Y)	Educational Int. Adenoma (N)	Control Adenoma (Y)	Control Adenoma (N)	Rate ratio with 95% CI	Weight (%)
Berger 2017	356	493	368	745	1.27 [1.13, 1.42]	10.93
Coe 2013	243	277	216	386	1.30 [1.13, 1.50]	9.11
Corley 2019	7815	13 082	3864	8402	1.19 [1.15, 1.23]	17.40
Evans 2019	300	550	265	568	1.11 [0.97, 1.27]	9.51
Hall 2010	140	273	121	429	1.54 [1.25, 1.90]	5.75
Kaminski 2016	2548	8067	2625	11 639	1.30 [1.24, 1.37]	16.34
Keswani 2015	1419	2220	684	1760	1.39 [1.29, 1.50]	14.18
Wallace 2017	3643	5030	2319	5161	1.35 [1.30, 1.41]	16.78
Overall (*)					1.29 [1.22, 1.37]	



Heterogeneity: $\tau^2 = 0.01$, $I^2 = 82.96\,\%$, $H^2 = 5.87$

Test of $\Theta_i = \Theta_j$: $Q(7) = 41.08$, $P = 0.00$

Test of $\Theta = 0$: $z = 8.36$, $P = 0.00$

Random-effects DerSimonian-Laird model

(*) Predicted interval [1.09, 1.53]

▶ **Fig. 1** Forest plot comparing the primary outcome of adenoma detection rate (ADR) pre-educational intervention and post-educational intervention. CI, confidence interval.

the three RCTs were analyzed using experimental and control groups, the rate ratio was still significant (RR 1.18, 95 % CI 1.06 to 1.31), but less pronounced than when these studies were analyzed using pre-intervention and post-intervention groups in order to compare them with other observational studies (RR 1.33, 95 % CI 1.29 to 1.37). Subgroup analyses are summarized in ▶ **Table 4** and are provided in detail in **Supplementary Fig. 3**.

Other sensitivity analyses and publication bias

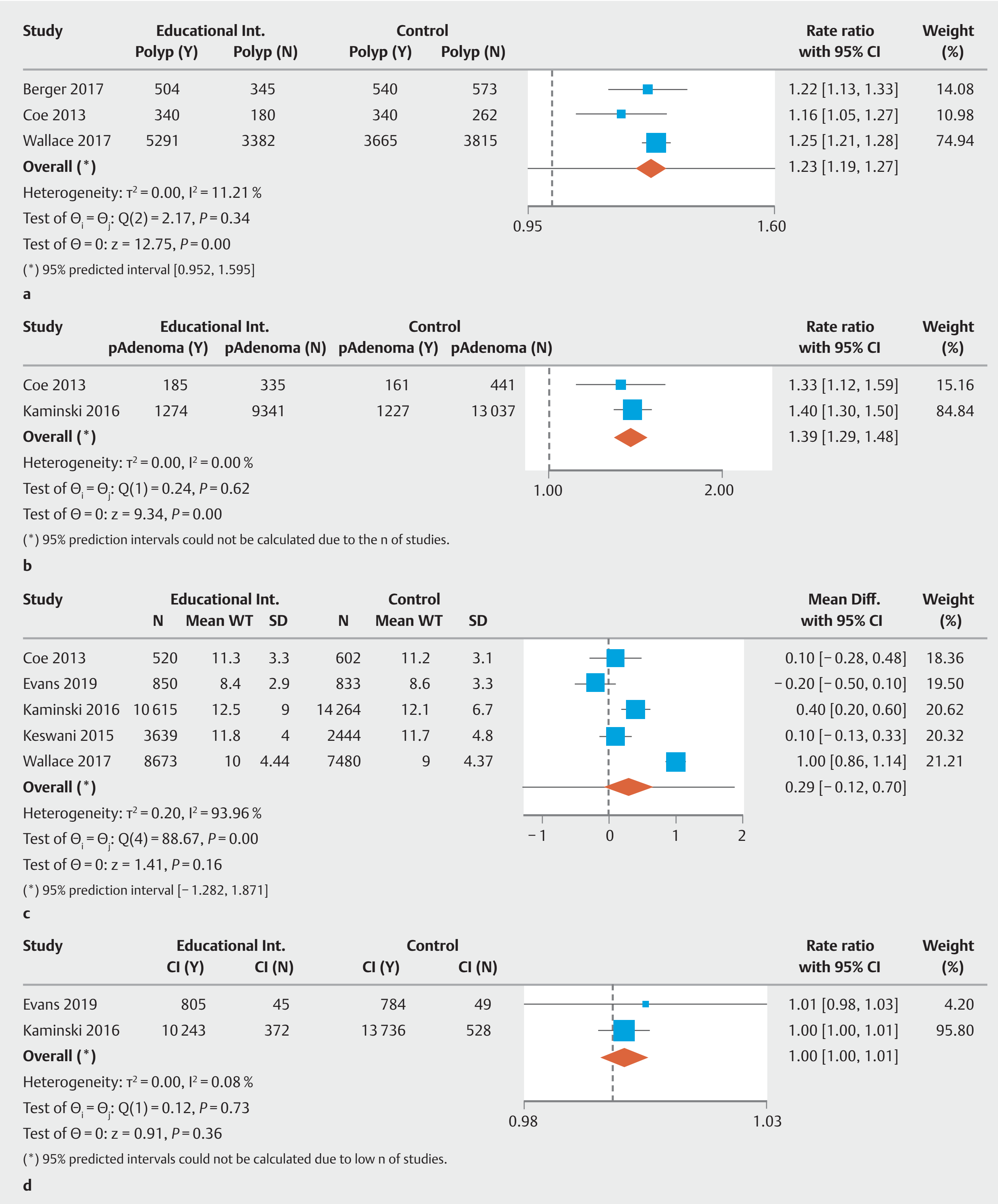
The findings for our primary outcome of ADR were robust to sensitivity analysis, as the RR did not change appreciably with exclusion of each study in turn or with analysis using a fixed effects model. There was no evidence of publication bias for the primary outcome by Egger's or Begg's tests, or by visual inspection of the funnel plot (**Supplementary Fig. 4**).

Discussion

In our meta-analysis of eight studies including 86,008 colonoscopies, educational interventions were associated with a significant 29 % relative increase in ADR. We also found that educational interventions are associated with significant improvements in PDR and pADR. Other quality indicators such as WT and CIR remained unchanged after educational interventions, though there was a trend toward increases in WT. Our results suggest that educational interventions aimed at independently practicing endoscopists contribute meaningfully to improvements in colonoscopy indicators.

ADR is the most widely endorsed colonoscopy quality indicator given its established inverse relationship with PCCRC [7] and CRC-related death [37]. Various strategies to optimize ADR specifically aimed at the endoscopist have been studied, including formalized audit and feedback [38], video-based assessments [39, 40], and combinations of these strategies [18], with mixed results. Short educational interventions targeting endoscopists, the subject of this review, have also been studied. Though these interventions vary considerably, they have in common the ultimate goal of improving the quality of colonoscopy performed by independent non-trainee endoscopists. EQUIP training, for example, is comprised of didactic presentations focusing on optimal withdrawal techniques and image recognition of neoplastic versus non-neoplastic polyps [27]. Conversely, skills enhancement and 'train-the-endoscopy-trainer' programs take a more holistic approach, providing a range of theoretical and hands-on training sessions, from basic to advanced [41]. These sessions are primarily focused on navigating the transition from unconscious competence to conscious competence [42], but also have the concurrent effect of optimizing technique and enabling more robust teaching, ultimately improving quality indicators [29]. Our findings confirm that these interventions have a demonstrable effect on several important colonoscopy quality indicators. Interestingly, studies assessing hands-on training interventions improved ADR at approximately the same rate as didactic interventions in our study. While hands-on training, including simulation, is currently recommended as part of endoscopy training curricula [43], its use among trained endoscopists has been relatively poorly studied.

Another interesting finding is that educational interventions were associated with significantly increased proximal adenoma detection rates. While the mechanisms for this change are not clearly explained by our results, several endoscopist-related factors could have been influenced by educational interventions, including technical skills, adequate air insufflation, washing and suctioning of debris and fluid, attention to flexures and folds, repetitive segment viewing, and torqueing maneuvers to enhance visualization. This finding is particularly important,



▶ **Fig. 2** Forest plot comparing **a** polyp detection rate (PDR), **b** proximal adenoma detection rate (pADR), **c** withdrawal time (WT), and **d** cecal intubation rate (CIR) pre-educational intervention and post-educational intervention. CI, confidence interval.

▶ **Table 4** Summary of subgroup analyses.

Subgroups	Pooled RR (95 % CI)	Inter-study heterogeneity (I^2)
No lag time* [28, 33, 34]	1.34 (1.28, 1.41)	11 %
Any lag time specified* [27, 29–32]	1.28 (1.19, 1.38)	82 %
Primarily screening colonoscopy [28–31, 33]	1.32 (1.25, 1.39)	60 %
Mixed populations [27, 32, 34]	1.21 (1.12, 1.30)	37 %
EQUIP-based studies [27, 28, 33]	1.34 (1.29, 1.39)	0 %
Non-EQUIP-based studies [29–32, 34]	1.28 (1.18, 1.39)	86 %
Hands-on skills training [29, 34]	1.22 (1.04, 1.43)	80 %
Didactic training [27, 28, 30–33]	1.32 (1.22, 1.42)	86 %
Published manuscripts [27–30, 34]	1.32 (1.25, 1.38)	59 %
Conference abstracts [31–33]	1.28 (1.14, 1.43)	71 %
RCT (pre-/post-) [27–29]	1.33 (1.29, 1.37)	0 %
RCT (control/intervention) [27–29]	1.18 (1.06, 1.31)	82 %
Observational studies [30–34]	1.27 (1.16, 1.40)	82 %
Multi-center studies [28, 29, 32]	1.28 (1.17, 1.39)	93 %
Single-center studies [27, 30, 31, 33, 34]	1.31 (1.19, 1.43)	64 %

RR, rate ratio of ADR post- versus preintervention; CI, confidence interval; RCT, randomized controlled trial; EQUIP [27], endoscopic quality improvement program.
* Lag time refers to time between intervention and start of post-intervention measurement of outcome(s).

given that missed proximal adenomas play an important role in the development of PCCRC [44].

While our results are encouraging, one must interpret them with a degree of caution. When all studies (including RCTs using pre-intervention and post-intervention groups) were analyzed as having quasi-experimental designs, there were significant improvements in ADR. When the RCTs were considered separately using experimental and control groups, the improvement in ADR was less pronounced. This could lead one to hypothesize a potential contribution from the Hawthorne effect to our overall pooled results, whereby the behavior of study subjects has the ability to change merely from their knowledge of being surveilled [45]. Potential pitfalls of educational interventions also need close consideration. Expert facilitators and instructors are often required to deliver such interventions, and they are often not readily available. Furthermore, considerable preparation and considerable resources are required to successfully conduct such educational programs. The cost to the system associated with these interventions thus needs to be considered, even if one can ultimately make a strong argument for overall cost savings through quality improvement and reduction of disease burden. In addition, it is unknown whether there are differences in the degree of improvement when an endoscopist independently seeks out additional training, versus when it is mandated. Finally, the durability of any improvements in quality seen as a result of educational interventions is poorly established, given that the length of post-intervention follow-up was not uniformly reported across studies. Kaminski *et al*. included a long post-intervention phase of 18 months and found that ADR continued to be higher than the preintervention period, but less pronounced compared to the first 6 months post-intervention [29]. Conversely, the long-term sustainability of ADR improvement with the EQUIP-3 intervention is less clear [28].

Of note, in a recent meta-analysis, we reported a significant association between endoscopist feedback and improvements in ADR, with a rate ratio of 1.21 [38]. The rate ratio of 1.29 we report in our present study reflects a potential added impact of educational interventions. However, all but one study included in our analysis employed endoscopist audit and feedback concurrently with educational interventions. As endoscopist feedback has previously been independently associated with improved colonoscopy quality [38], a potential confounding effect needs to be considered. The lone study assessing educational interventions alone (without feedback) in our analysis demonstrated non-statistically significant ADR improvements. Thus, further study is required to reliably determine the effect of educational interventions without any form of audit and feedback, and to assess the direct contributions of each of these measures on ADR and other quality indicators.

Our review has a number of strengths. A separate meta-analysis also recently evaluated the effect of educational interventions on ADR, but restricted inclusion to the 3 RCTs only, thereby missing several important studies assessing this question [46]. Furthermore, subgroup analyses were limited by the small number of input studies [46]. Our comprehensive search strategy included both randomized and non-randomized study designs as well as conference abstracts, thereby resulting in 13 studies included in our systematic review, with eight ultimately included in our quantitative meta-analysis. We also carried out several subgroup analyses to better understand potential sources of heterogeneity between studies. Importantly, we identified that studies restricted to primarily screening populations, the most common indication for colonoscopy, displayed considerably less inter-study heterogeneity. Lastly, we employed the GRADE approach to assess and summarize the certainty of the evidence leading to our conclusions.

Our study also has several limitations. For our overall analysis, we included only the interventional arms of RCTs and considered them as quasi-experimental studies in order to be able to compare them to the other included observational studies. However, to mitigate this, we also conducted a sensitivity analysis pooling only RCT data; the pooled interventional group still had ADR improvements compared to the control group. However, one should be cognizant that the degree of improvement seen with RCTs was somewhat lower than seen with quasi-experimental studies. This is in part owing to the fact that there was a (lesser) degree of ADR improvement in two of the RCT control groups as well. Thus, the overall pooled magnitude of ADR improvement should be interpreted with caution and in conjunction from the RCT-specific estimates. Secondly, there is a present lack of understanding of the mechanism(s) linking educational interventions with ADR improvement. This is important given that WT and CIR, whose improvements would be plausible mechanisms, remained unchanged. Thirdly, we opted to include gray literature. Although we believe this reduced publication bias, we acknowledge this may represent a limitation as it increases the heterogeneity due to incomplete details about the methodology and the lack of a peer review process. Encouragingly, in subgroup analyses, estimates of the primary outcome remained unchanged based on publication type. Finally, the majority of included studies were observational by design, and though attempts were made to report results adjusted for known confounders, the inability to adjust for unknown confounders must be acknowledged.

Conclusion

In conclusion, we found evidence that educational interventions improve ADR when conducted among independent endoscopists. Furthermore, they are also associated with improvements in pADR and overall PDR. Though the certainty of the evidence leading to conclusions on the primary outcome was low, we believe our findings are important. Future research should prioritize addressing important gaps, such as assessing the durability of the intervention, the impact on low- versus high-performers, and whether hands-on training or multimodal training are superior to didactic educational sessions.

Competing interests

The authors declare that they have no conflict of interest.

References

[1] American Cancer Society. Key Statistics for Colorectal Cancer. Atlanta, GA: American Cancer Society; 2020: Accessed: 2020 May 22 https://www.cancer.org/cancer/colon-rectal-cancer/about/key-statistics.html

[2] Vogelstein B, Fearon ER, Hamilton SR et al. Genetic alterations during colorectal-tumor development. N Engl J Med 1988; 319: 525–532

[3] Zauber AG, Winawer SJ, O'Brien MJ et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med 2012; 366: 687–696

[4] Ran T, Cheng CY, Misselwitz B et al. Cost-effectiveness of colorectal cancer screening strategies-a systematic review. Clin Gastroenterol Hepatol 2019; 17: 1969–1981.e1915

[5] Rex DK, Schoenfeld PS, Cohen J et al. Quality indicators for colonoscopy. Am J Gastroenterol 2015; 110: 72–90

[6] Rees CJ, Gibson TS, Rutter MD et al. UK key performance indicators and quality assurance standards for colonoscopy. Gut 2016; 65: 1923–1929

[7] Kaminski MF, Regula J, Kraszewska E et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med 2010; 362: 1795–1803

[8] Singh S, Singh PP, Murad MH et al. Prevalence, risk factors, and outcomes of interval colorectal cancers: a systematic review and meta-analysis. Am J Gastroenterol 2014; 109: 1375–1389

[9] Morris EJ, Rutter MD, Finan PJ et al. Post-colonoscopy colorectal cancer (PCCRC) rates vary considerably depending on the method used to calculate them: a retrospective observational population-based study of PCCRC in the English National Health Service. Gut 2015; 64: 1248–1256

[10] Kaminski MF, Gibson TS, Bugajski M et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. Endoscopy 2017; 49: 378–397

[11] Hilsden RJ, Rose SM, Dube C et al. Defining and applying locally relevant benchmarks for the adenoma detection rate. Am J Gastroenterol 2019; 114: 1315–1321

[12] Mehrotra A, Morris M, Gourevitch RA et al. Physician characteristics associated with higher adenoma detection rate. Gastrointest Endosc 2018; 87: 778–786.e775

[13] Ezaz G, Leffler DA, Beach S et al. Association between endoscopist personality and rate of adenoma detection. Clin Gastroenterol Hepatol 2019; 17: 1571–1579.e1577

[14] Sanaka MR, Gohel T, Podugu A et al. Adenoma and sessile serrated polyp detection rates: variation by patient sex and colonic segment but not specialty of the endoscopist. Dis Colon Rectum 2014; 57: 1113–1119

[15] Heresbach D, Barrioz T, Lapalus MG et al. Miss rate for colorectal neoplastic polyps: a prospective multicenter study of back-to-back video colonoscopies. Endoscopy 2008; 40: 284–290

[16] Facciorusso A, Triantafyllou K, Murad MH et al. Compared Abilities of endoscopic techniques to increase colon adenoma detection rates: a network meta-analysis. Clin Gastroenterol Hepatol 2019; 17: 2439–2454.e2425

[17] Forbes N, Mohamed R, Raman M. Learning curve for endoscopy training: Is it all about numbers? Best Pract Res Clin Gastroenterol 2016; 30: 349–356

[18] Shaukat A, Oancea C, Bond JH et al. Variation in detection of adenomas and polyps by colonoscopy and change over time with a performance improvement program. Clin Gastroenterol Hepatol 2009; 7: 1335–1340

[19] Rajasekhar PT, Rees CJ, Bramble MG et al. A multicenter pragmatic study of an evidence-based intervention to improve adenoma detection: the Quality Improvement in Colonoscopy (QIC) study. Endoscopy 2015; 47: 217–224

[20] Moher D, Shamseer L, Clarke M et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015; 4: 1

[21] The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses. Ottawa, ON: The Ottawa Hospital; 2019: Accessed: 2020 May 22 http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

[22] Guyatt GH, Oxman AD, Vist GE et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008; 336: 924–926

[23] Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011; 342: d549

[24] Cochrane Handbook for Systematic Reviews of Interventions. The Cochrane Collaboration; 2011: Accessed: 2020 May 22 http://handbook-5-1.cochrane.org/front_page.htm

[25] Egger M, Smith DG, Schneider M et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997; 315: 629–634

[26] Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994; 50: 1088–1101

[27] Coe SG, Crook JE, Diehl NN et al. An endoscopic quality improvement program improves detection of colorectal adenomas. Am J Gastroenterol 2013; 108: 219–226

[28] Wallace MB, Crook JE, Thomas CS et al. Effect of an endoscopic quality improvement program on adenoma detection rates: a multicenter cluster-randomized controlled trial in a clinical practice setting (EQUIP-3). Gastroint Endosc 2017; 85: 538–545

[29] Kaminski MF, Anderson J, Valori R et al. Leadership training to improve adenoma detection rate in screening colonoscopy: A randomised trial. Gut 2016; 65: 616–624

[30] Keswani R, Yadlapati R, Gleason K et al. Physician report cards and implementing standards of practice are both significantly associated with improved screening colonoscopy quality. Am J Gastroenterol 2015; 110: 1134

[31] Hall B, Benson M, Pfau P et al. Improved Adenoma detection rates at an academic gastroenterology unit following department colonoscopy assessment. Gastroint Endosc 2010; 71: AB107–108

[32] Corley D, Jensen C, Lee J et al. Increasing physician adenoma detection rate is associated with a reduced risk of post-colonoscopy colorectal cancer. Gastroenterology 2019; 156: S151

[33] Berger D, Hubbard E, Greenspan M et al. Non-neoplastic resection rate and adenoma detection rate increase after intervention to improve and monitor ADR. Gastroint Endosc 2017; 85: AB395

[34] Evans B, Pace D, Borgaonkar M et al. Effect of an educational intervention on colonoscopy quality outcomes. Surg Endosc 2019: doi:10.1007/s00464-019-07304-w

[35] Rank J. Increased Adenoma Find Rate (AFR) and proportionately greater serrated adenoma find rate (safr) following a systematic quality improvement program in a large community gastroenterology practice. Gastroenterology 2011; 140: S199–200

[36] Salden B, Bouwens M, Winkens B et al. Systematic training improved the quality of colonoscopy by reducing the variation of quality indicators among providers. Gastroenterology 2012; 142: S2218

[37] Corley DA, Jensen CD, Marks AR et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 2014; 370: 1298–1306

[38] Bishay K, Causada-Calo N, Scaffidi MA et al. Endoscopist feedback is associated with improvements in colonoscopy quality indicators: a systematic review and meta-analysis. Gastroint Endosc 2020: doi:10.1016/j.gie.2020.03.3865 Online ahead of print.

[39] Madhoun MF, Tierney WM. The impact of video recording colonoscopy on adenoma detection rates. Gastroint Endosc 2012; 75: 127–133

[40] Scaffidi MA, Grover SC, Carnahan H et al. A prospective comparison of live and video-based assessments of colonoscopy performance. Gastroint Endosc 2018; 87: 766–775

[41] Waschke KA, Anderson J, Macintosh D et al. Training the gastrointestinal endoscopy trainer. Best Pract Res Clin Gastroenterol 2016; 30: 409–419

[42] Peyton J. Teaching and learning in medical practice. Rickmansworth, UK: Manticore Europe Limited; 1998

[43] Ekkelenkamp VE, Koch AD, de Man RA et al. Training and competence assessment in GI endoscopy: a systematic review. Gut 2016; 65: 607–615

[44] le Clercq CM, Bouwens MW, Rondagh EJ et al. Postcolonoscopy colorectal cancers are preventable: a population-based study. Gut 2014; 63: 957–963

[45] Sedgwick P, Greenwood N. Understanding the Hawthorne effect. BMJ 2015; 351: h4672

[46] Lim S, Hammond S, Park J et al. Training interventions to improve adenoma detection rates during colonoscopy: a systematic review and meta-analysis. Surg Endosc 2019: doi:10.1007/s00464-019-07153-7