Accepted Manuscript online: 2022-12-07 Article published online: 2023-01-26



# Colonoscopy quality improvement after initial training: A cross-sectional study of intensive short-term training

OPEN ACCESS

Authors
Anna Lisa Schult[1,2], Geir Hoff[1,3,4], Øyvind Holme[1,5,6], Edoardo Botteri[1,7], Birgitte Seip[1,8], Kristin Ranheim Randel[1], Ole Darre-Næss[2], Tanja Owen[9], Jens Aksel Nilsen[2], Dung Hong Nguyen[9], Kristin Johansen[1], Thomas de Lange[10,11,12]

Institutions
1 Section for Colorectal Cancer Screening, Cancer Registry of Norway, Oslo, Norway
2 Department of Medicine, Vestre Viken Hospital Trust Bærum, Gjettum, Norway
3 Institute of Clinical Medicine, University of Oslo, Oslo, Norway
4 Department of Research and Development, Telemark Hospital Trust, Skien, Norway
5 Institute of Health and Society, University of Oslo, Oslo, Norway
6 Department of Medicine, Sørlandet Hospital Trust, Kristiansand, Norway
7 Department of Research, Cancer Registry of Norway, Oslo, Norway
8 Department of Medicine, Vestfold Hospital, Tønsberg, Norway
9 Department of Medicine, Østfold Hospital Trust, Grålum, Norway
10 Department of Medicine and Emergencies Sahlgrenska University Hospital-Mölndal, Region Västra Götaland, Sweden
11 Department of Molecular and Clinical Medicine, Sahlgrenska Academy, University of Gothenburg, Sweden
12 Department of Medical Research, Vestre Viken Hospital Trust Bærum, Gjettum, Norway

submitted 24.6.2022
accepted after revision 6.12.2022
published online 7.12.2022

Bibliography
Endosc Int Open 2023; 11: E117–E127
DOI 10.1055/a-1994-6084
ISSN 2364-3722

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

Supplementary material is available under https://doi.org/10.1055/a-1994-6084

Corresponding author
Anna Lisa Schult, Cancer Registry of Norway, P.O. Box 5313, 0304 Oslo, Norway
Phone: +004741393003
Fax: +004741393003
anna.schult@kreftregisteret.no

## ABSTRACT

**Background and study aims** High-quality is crucial for the effectiveness of colonoscopy and can be achieved by high-quality training and verified with assessment of key performance indicators (KPIs) for colonoscopy such as cecum intubation rate (CIR), adenoma detection rate (ADR) and adequate polyp resection. Typically, trainees achieve adequate CIR after 275 procedures, but little is known about learning curves for KPIs after initial training.

**Methods** This cross-sectional study includes work-up colonoscopies after a positive screening test with fecal occult blood testing (FIT) or sigmoidoscopy, performed by either trainees after 300 training colonoscopies or by consultants. Outcome measures were KPIs. We assessed inter-endoscopist variation in trainees and learning curves for trainees as a group. We also compared KPIs for trainees and consultants as a group.

**Results** Data from 6,655 colonoscopies performed by 21 trainees and 921 colonoscopies performed by 17 consultants were included. Most trainees achieved target standards for main KPIs. With time, trainees shortened cecum intubation time and withdrawal time without decreasing their ADR, reduced the proportion of painful colonoscopies, and increased the adequate polyp resection rate (all $P<0.01$). Compared to consultants, trainees had higher CIR (97.7% vs. 96.3%, $P=0.02$), ADR after positive FIT (57.6% vs. 50.3%, $P<0.01$), and proximal ADR after sigmoidoscopy screening (41.1% vs. 29.8%; $P<0.01$), higher adequate polyp resection rate (94.9% vs. 93.1%, $P=0.01$) and fewer serious adverse events (0.65% vs. 1.41%, $P=0.02$).

**Conclusions** Trainees performed high-quality colonoscopies and achieved international target standards. Several KPIs continuously improved after initial training. Trainees outperformed consultants on several KPIs.

## Introduction

Colonoscopy is the cornerstone in colorectal cancer (CRC) screening, investigation of bowel symptoms and surveillance for polyps and inflammatory bowel disease.

Several societies have published guidelines for colonoscopy quality [1–3] measured by key performance indicators (KPIs). KPIs vary widely among endoscopists, and inadequate performance jeopardizes patient outcome [4]. High cecum intubation rate (CIR), high adenoma detection rate (ADR) and adequate polyp resection techniques are crucial to reduce post-colonoscopy CRCs (PCCRCs) [5–7]. Additionally, short withdrawal time is associated with low ADR [8]. Patient reported outcome and experience measures are also important, since poor patient experiences represent barriers to colonoscopy attendance, although no target standard exists [1]. Endoscopists who have a high CIR cause less discomfort and use less sedation [9]. Finally, serious adverse events (SAEs) like perforation and bleeding are rare in high-quality colonoscopies [1].

High-quality training is required to achieve target KPIs, but current guidelines focus more on numbers and recommend at least 275 supervised colonoscopies before colonoscopies are performed independently [10, 11]. CIR is considered the most important measure of colonoscopy competence [10, 11]. However, studies have demonstrated wide variation in learning curves for various KPIs [12–15]. Consequently, it may be more appropriate to focus on the achievement of a wider range of target KPIs [16, 17]. Little is known about improvement of KPIs after initial colonoscopy training and whether learning curves are different for various KPIs.

This study aimed to assess individual colonoscopy performance in trainees, whether there is a continuous improvement of specific KPIs after initial colonoscopy training, and whether it is feasible to schedule repetitive assessments of their competency. It also compared endoscopy trainees (named trainees) and gastroenterology consultants (named consultants), serving as a reference for adequate endoscopy performance.

## Methods

### Design

This cross-sectional study assessed colonoscopy KPIs for trainees and consultants within a Norwegian population-based, randomized CRC screening trial. The trainees and consultants were employed at the section of gastroenterology at two secondary care hospitals housing their respective screening center, and the study was coordinated by the Cancer Registry of Norway. The screening trial invited 140.000 persons aged 50–74 to once-only sigmoidoscopy or fecal immunochemical test (FIT) every other year, described in detail elsewhere [18]. We included data from work-up colonoscopies after a positive primary screening test from April 2012 until April 2020. The screening trial was approved by the Regional Committee for Medical Research Ethics in Southeast Norway (2011/1272) and is registered at clinicaltrials.gov (NCT 01538550).

### Consultants and trainees

Prior to the screening trial, consultants at both screening centers had attended Joint Advisory Group (JAG) on Endoscopy's train-the-colonoscopy-trainer (TCT) courses [19]. TCT courses focus on competency-based training and aim to provide endoscopists with conscious competence in colonoscopy. All consultants performing colonoscopies included in this trial were specialists in gastroenterology and their KPIs serve as a reference value for performance comparison. Colonoscopies performed by gastroenterology surgeons were not included because surgeons only conducted a very limited number of colonoscopies.

The trainees at both screening centers were specifically recruited to perform sigmoidoscopies and colonoscopies in the screening trial and had no or very limited previous endoscopy experience. The first three to six months of their employment, most of them were exclusively trained in endoscopy by one-to-one supervision. Delegates from TCT courses were in charge of the training and supervision of the trainees, but also consultants not specifically trained in TCT courses supervised the trainees.

After initial training, trainees worked half time as endoscopists in the screening trial and half time as trainees in internal medicine or gastroenterology.

During the screening trial, trainees had access to their own continuously updated KPIs. Staff meetings were held twice a year to assess and provide feedback on the individual trainee's performance and align endoscopy routines.

### Screening participants and colonoscopies

Screening participants with a positive test (FIT threshold > 15 µg hemoglobin/g feces or sigmoidoscopy with any polyp sized ≥ 10 mm, advanced adenoma, or ≥ 3 adenoma), were referred to work-up colonoscopy [18]. In FIT positives, it was registered whether the test was performed in the first, second or third round of testing. Split dose bowel preparation was recommended, and $CO_2$ was used for insufflation. The Olympus Exera II/III systems (Olympus H180DL/I, CF-HQ190 L/I, PCF-PH190 L/I, PCF-H190DL/I) were used for all colonoscopies. Water-assisted intubation was performed at the endoscopist's discretion and magnetic endoscopic imaging (MEI) was available for all colonoscopies. The endoscopists assessed bowel cleansing quality with a four-point rating scale (good, acceptable, partially poor, or poor) [18]. Sedation or analgesia was provided on demand.

## Data collection and outcome measures

Colonoscopies performed by trainees who had finished a training program including 300 colonoscopies and who had performed at least 30 colonoscopies within the screening trial after the training period, were included as well as all colonoscopies performed by consultants. Data from Gastronet, the Norwegian national quality assurance program for endoscopy, were used to determine the date for the 300$^{th}$ colonoscopy. For some trainees, only the month of colonoscopy performance was recorded. In these cases, all colonoscopies from the month following the one with the 300$^{th}$ colonoscopy were included in the study.

Cecum intubation was defined as passing the ileocecal valve and visualizing the entire cecum and CIR was based on each screening participant's first colonoscopy, since cecum intubation was not intended in many additional colonoscopies required to achieve polyp-free colon. Cecum intubation time was defined as the time spent advancing the endoscope from anus to cecum and was measured for all complete colonoscopies.

For all lesions detected at colonoscopy, the appearance (pedunculated or sessile) and size of removed lesions as well as the resection technique (biopsy forceps or snare polypectomy) was recorded.

The World Health organization guidelines for histological assessment of neoplastic lesions was followed [20]. Advanced adenoma was defined as adenoma ≥ 10 mm, villous components of at least 25 %, and/or high-grade dysplasia. Adenoma detection rate (ADR) and advanced ADR (AADR) were calculated only for colonoscopies after positive FIT, since distal lesions discovered during primary sigmoidoscopy screening were removed during the initial examination and another endoscopist often performed the colonoscopy. For colonoscopies following a positive sigmoidoscopy, ADR and AADR proximal to and including the splenic flexure were calculated. Detection rates were calculated as the proportion of participants with at least one lesion detected at their first colonoscopy among the total number of first colonoscopies.

Withdrawal time was defined as total procedure time minus intubation time when no biopsy or polypectomy was performed [1].

Adequate polyp resection technique was defined as snare polypectomy of polyps (traditional serrated adenoma, hyperplastic polyp, sessile serrated lesion, non-advanced adenoma, and advanced adenoma) larger than 3 mm [1]. Adequate polyp resection rate was the proportion of adequately resected lesions.

The Gastronet questionnaire was used to record patients' satisfaction (not, a little, moderately, or very satisfied) and procedural pain during colonoscopy (none, slight, moderate, or severe pain) on a 4-point Likert scale [21]. The participants received the questionnaire at the screening center and were asked to return it in a prepaid return envelope the following day. Pain was dichotomized into painful (moderate or severe pain) versus non-painful (slight or no pain) colonoscopy. Due to missing identification of participants, data on patient satisfaction and patient-reported pain from 2013 were excluded. Data on patient satisfaction from 2012 were excluded due to the use of a 2-point Likert scale. Rates were calculated among responders.

SAEs limited to perforations, significant bleedings, and death occurring within 30 days after colonoscopy were assessed by scrutinizing the health trusts' electronic medical report system in case of polypectomy of polyps larger than one cm or a re-contact coded as bleeding or perforation. Significant bleeding was defined as bleeding causing hospitalization, blood transfusion, repeated endoscopy, radiological intervention, or surgery. Perforation was defined as detection of free intraperitoneal air on computed tomography. Mortality within 30 days after endoscopy was assessed by linkage to the Norwegian population registry and a possible relation to endoscopy was evaluated by the study personnel scrutinizing the medical record.

Outcome measures were calculated for the individual trainee and for the comparison of trainees and consultants as groups.

## Statistics

Continuous variables were expressed as means and interquartile ranges (IQR) or 95 % confidence intervals (CI); binary variables were expressed as numbers and percentages. At univariate analysis, we used Chi-squared test and Fisher exact test to compare binary KPIs. For continuous KPIs, we used one-way ANOVA to compare the individual trainees and Student t-test to compare KPIs between trainees and consultants as groups.

At multivariable analysis, we fitted logistic (for binary KPIs) and linear (for continuous KPIs) regression models to compare trainees and consultants as groups. To test for the effect of the individual trainees on different KPIs, we compared nested regression models with and without the individual trainees through the likelihood ratio test.

Age and sex of the screenees were included in all models. Further covariates included in regression models depended on the KPI:

1. ADRs and AADRs in FIT screening were additionally adjusted for FIT rounds due to decreasing detection rates in subsequent FIT rounds.
2. SAEs, pain, and satisfaction were additionally adjusted for polyp size and appearance which may impact the outcome. SAEs were also adjusted for the use of anticoagulants or antiplatelet therapy.

For comparison of detection rates between individual trainees, those without lesion detection were excluded from analyses. Only colonoscopies with polypectomy were included in multivariable logistic regression analyses comparing SAEs between trainees and consultants as a group.

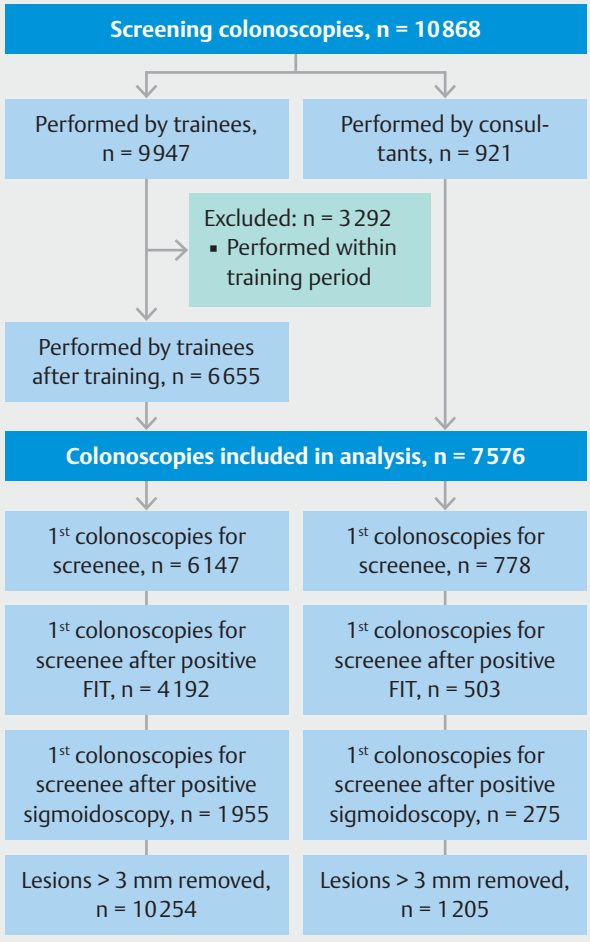
Learning curves were evaluated per year of screening. The first year was 365 days after completion of the training and accordingly the following years were 365 days each. To assess learning curves, we used logistic (for binary KPIs) and linear (for continuous KPIs) regression models with time in days as the continuous independent variable. Models were adjusted for screenees' age and sex. Models assessing the proportion of painful colonoscopies and patient satisfaction were additionally adjusted for polyp size and appearance.

*P* values and odds ratios (ORs) with 95 % confidence intervals (CIs) were reported. All tests were two-sided, and $P<0.05$ was considered statistically significant.

Statistical analyses were performed using Stata statistical software version 17.0 (StataCorp, College Station, Texas, United States).

## Results

In total 7,576 colonoscopies were included in the analysis (▶ **Fig. 1**). After training, trainees performed 6.655 screening colonoscopies, and consultants performed 921 colonoscopies. Overall, 24 trainees performed colonoscopies in the screening trial. Of these, 21 were included in the present study since they performed at least 30 colonoscopies (range 65–1,140) after the initial 300 training colonoscopies (▶ **Table 1**). In addition, 17 consultants performing colonoscopies in the trial were included.




▶ **Fig. 1** Flowchart.

▶ **Table 1** Baseline colonosopy data for trainees after training.

| Trainee | Number of colonoscopies | Age of screenees, years, median (IQR)* | Female sex screenees, % |
|---|---|---|---|
| 1 | 1,140 | 65.9 (60.3–70.5) | 43.1 |
| 2 | 642 | 66.9 (60.6–71.3) | 46.4 |
| 3 | 538 | 67.3 (61.3–71.3) | 41.3 |
| 4 | 514 | 66.4 (61.8–71.3) | 42.6 |
| 5 | 482 | 66.8 (61.9–71.8) | 44.0 |
| 6 | 414 | 65.6 (59.9–71.5) | 44.7 |
| 7 | 392 | 65.0 (59.4–69.5) | 41.8 |
| 8 | 333 | 66.1 (60.1–70.3) | 42.6 |
| 9 | 325 | 65.8 (60.3–71.1) | 40.3 |
| 10 | 306 | 64.1 (57.2–69.7) | 40.8 |
| 11 | 277 | 65.6 (59.3–70.4) | 42.2 |
| 12 | 243 | 64.0 (57.4–69.6) | 36.6 |
| 13 | 229 | 66.1 (60.8–71.6) | 40.6 |
| 14 | 177 | 66.6 (61.6–71.3) | 40.1 |
| 15 | 113 | 67.6 (61.8–72.1) | 46.9 |
| 16 | 112 | 65.6 (60.0–70.4) | 33.9 |
| 17 | 99 | 64.3 (58.8–69.0) | 48.5 |
| 18 | 92 | 66.0 (60.0–72.3) | 27.2 |
| 19 | 83 | 65.0 (59.3–70.0) | 48.2 |
| 20 | 79 | 66.2 (61.0–71.1) | 34.2 |
| 21 | 65 | 66.4 (62.4–71.5) | 35.4 |

IQR, interquartile range.
* Age either at invitation to sigmoidoscopy or at invitation to fecal immunochemical test, which was positive and followed by subsequent colonoscopy.

### Inter-endoscopist variation in individual trainees

The overall CIR was 97.7 % (range between endoscopists 94.2 %–100 %) and it differed significantly both at univariate ($P=0.01$) and multivariable analysis ($P=0.01$) (▶ **Table 2**, ▶ **Table 3**, ▶ **Fig. 2**). Mean cecum intubation time for the individual trainee varied between 8.7 min and 17.7 min ($P<0.01$ at univariate and multivariable analysis) and overall mean cecum intubation time was 13.7 min (95 % confidence interval [CI] 13.4 min-13.9 min). ADR at colonoscopy after a positive FIT screening was 57.6 % and AADR was 24.9 % (▶ **Table 3**). The inter-endoscopist variation in ADR ranged from 48.9 %- 69.5 % ($P=0.01$) and AADR from 13.6 %- 43.9 % ($P<0.01$) after a positive FIT (▶ **Table 2**, ▶ **Fig. 1**). At multivariable analysis, the inter-endoscopist variation in ADR was confirmed ($P=0.02$), while there was no difference in AADR ($P=0.16$). In colonoscopies after a positive sigmoidoscopy screening, the overall proximal ADR was 41.1 % and proximal AADR 8.6 %. Similar to the FIT-positive group, there was a significant inter-endoscopist variation

▶ **Table 2** Comparison of colonoscopy key performance indicators among trainees.

| Trainee | CIR, n/N (%)[1] | ADR after FIT, n/N (%)[1] | AADR after FIT, n/N (%)[1] | Prox. ADR after sigmoidoscopy, n/N (%)[1] | Prox. AADR after sigmoidoscopy, n/N (%)[1] | Cecum intubation time, mean (95 % CI), min | Withdrawal time, mean (95 % CI) min | Adequate resection technique, n/N (%)[2] | Painful colonoscopy, n/N (%)[3] | Very satisfied, n/N (%)[4] |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1,024/1,038 (98.7) | 363/685 (53.0) | 182/685 (26.6) | 124/353 (35.1) | 36/353 (10.2) | 14.3 (13.7–14.9) | 9.9 (9.4–10.3) | 1,465/1,519 (96.4) | 102/670 (15.2) | 577/623 (92.6) |
| 2 | 586/600 (97.7) | 254/426 (59.6) | 80/426 (18.8) | 82/174 (47.1) | 7/174 (4.0) | 15.0 (14.1–15.9) | 13.2 (12.0–14.3) | 905/943 (96.0) | 105/460 (22.8) | 415/459 (90.4) |
| 3 | 493/504 (97.8) | 209/373 (56.0) | 77/373 (20.6) | 55/131 (42.0) | 15/131 (11.5) | 11.6 (10.8–12.3) | 13.5 (12.6–14.5) | 642/665 (96.5) | 54/391 (13.8) | 369/392 (94.1) |
| 4 | 486/489 (99.4) | 194/331 (58.6) | 68/331 (20.5) | 83/158 (52.5) | 13/158 (8.2) | 13.5 (12.8–14.3) | 15.5 (14.2–16.8) | 1,118/1,210 (92.4) | 85/372 (22.9) | 336/371 (90.6) |
| 5 | 454/467 (97.2) | 195/374 (52.1) | 66/374 (17.7) | 29/93 (31.2) | 3/93 (3.2) | 13.9 (13.2–14.5) | 10.1 (9.5–10.7) | 538/574 (93.7) | 73/346 (21.1) | 319/346 (92.2) |
| 6 | 367/380 (96.6) | 147/257 (57.2) | 64/257 (24.9) | 59/123 (48.0) | 9/123 (7.3) | 15.3 (14.4–16.3) | 14.0 (12.7–15.3) | 579/641 (90.3) | 81/310 (26.1) | 286/312 (91.7) |
| 7 | 335/350 (95.7) | 129/204 (63.2) | 70/204 (34.3) | 59/146 (40.4) | 14/146 (9.6) | 15.1 (14.0–16.2) | 11.5 (10.5–12.5) | 655/692 (94.7) | 18/146 (12.3) | 117/130 (90.0) |
| 8 | 292/301 (97.0) | 139/214 (65.0) | 66/214 (30.8) | 42/87 (48.3) | 7/87 (8.1) | 11.2 (10.3–12.0) | 15.7 (14.0–17. 3) | 591/608 (97.2) | 59/265 (22.3) | 235/264 (89.0) |
| 9 | 284/293 (96.9) | 95/176 (54.0) | 44/176 (25.0) | 35/117 (29.9) | 9/117 (7.7) | 8.9 (8.2–9.6) | 10.7 (9.9–11.6) | 432/436 (99.1) | 75/221 (33.9) | 194/219 (88.6) |
| 10 | 281/283 (99.3) | 111/164 (67.7) | 55/164 (33.5) | 55/119 (46.2) | 13/119 (10.9) | 15.3 (14.2–16.5) | 13.5 (11.9–15.1) | 616/642 (96.0) | 29/171 (17.0) | 150/162 (92.6) |
| 11 | 252/259 (97.3) | 105/180 (58.3) | 51/180 (28.3) | 25/79 (31.6) | 1/79 (1.3) | 12.8 (11.5–14.0) | 14.1 (12.4–15.8) | 355/357 (99.4) | 33/198 (16.7) | 131/138 (94.9) |
| 12 | 221/224 (98.7) | 85/137 (62.0) | 53/137 (38.7) | 37/87 (42.5) | 9/87 (10.3) | 14.4 (13.2–15.7) | 14.5 (12.8–16.2) | 417/427 (97.7) | 17/89 (19.1) | 71/77 (92.2) |
| 13 | 200/206 (97.1) | 74/143 (51.7) | 30/143 (21.0) | 29/63 (46.0) | 8/63 (12.7) | 14.0 (12.9–15.0) | 15.8 (14.4–17.1) | 299/312 (95.8) | 44/166 (26.5) | 159/167 (95.2) |
| 14 | 156/161 (96.9) | 49/82 (59.8) | 22/82 (26.8) | 32/79 (40.5) | 7/79 (8.9) | 8.7 (7.4–9.9) | 25.1 (18.2–32.0) | 331/374 (88.5) | 49/125 (39.2) | 115/126 (91.3) |
| 15 | 99/102 (97.1) | 58/98 (59.2) | 21/98 (21.4) | 0/4 (0.0) | 0/4 (0.0) | 17.7 (15.6–19.8) | 13.4 (11.3–15.4) | 128/140 (91.4) | 28/72 (38.9) | 70/74 (94.6) |
| 16 | 97/102 (95.1) | 42/77 (54.5) | 21/77 (27.3) | 10/25 (40.0) | 6/25 (24.0) | 10.3 (8.9–11.8) | 9.6 (7.7–11.5) | 119/119 (100.0) | 21/83 (25.3) | 78/82 (95.1) |

▶ **Table 2** (Continuation)

| Trainee | CIR, n/N (%)[1] | ADR after FIT, n/N (%)[1] | AADR after FIT, n/N (%)[1] | Prox. ADR after sigmoidoscopy, n/N (%)[1] | Prox. AADR after sigmoidoscopy, n/N (%)[1] | Cecum intubation time, mean (95 % CI), min | Withdrawal time, mean (95 % CI) min | Adequate resection technique, n/N (%)[2] | Painful colonoscopy, n/N (%)[3] | Very satisfied, n/N (%)[4] |
|---|---|---|---|---|---|---|---|---|---|---|
| 17 | 88/93 (94.6) | 37/57 (64.9) | 25/57 (43.9) | 10/36 (27.8) | 4/36 (11.1) | 14.1 (12.1–16.2) | 14.4 (11.6–17.2) | 162/177 (91.5) | 15/61 (24.6) | 57/60 (95.0) |
| 18 | 81/86 (94.2) | 33/59 (55.9) | 10/59 (17.0) | 14/27 (51.9) | 3/27 (11.1) | 17.4 (15.3–19.6) | 18.8 (14.2–23.3) | 102/103 (99.0) | 11/55 (20.0) | 50/55 (90.9) |
| 19 | 77/77 (100.0) | 31/49 (63.3) | 17/49 (34.7) | 14/28 (50.0) | 3/28 (10.7) | 15.6 (12.8–18.4) | 15.9 (5.4–26.3) | 136/168 (81.0) | 9/51 (17.7) | 40/43 (93.0) |
| 20 | 72/73 (98.6) | 23/47 (48.9) | 12/47 (25.5) | 10/26 (38.5) | 2/26 (7.7) | 15.1 (12.6–17.6) | 22.1 (17. 9–26.2) | 77/77 (100.0) | 6/61 (9.8) | 59/62 (95.2) |
| 21 | 58/59 (98.3) | 41/59 (69.5) | 8/59 (13.6) | 0/0 (–) | NA (–) | 17.7 (14.8–20.5) | 21.8 (17.8–25.9) | 67/70 (95.7) | 10/40 (25.0) | 38/40 (95.0) |
| *P* (univariate) | 0.01 | 0.01 | <0.01 | <0.01 | 0.09 | <0.01 | <0.01 | <0.01 | <0.01 | 0.06 |
| *P* (multivariate) | 0.01 | 0.02 | 0.16 | <0.01 | 0.03 | <0.01 | <0.01 | <0.01 | <0.01 | 0.43 |

CIR, cecum intubation time; ADR, adenoma detection rate; FIT, fecal immunochemical test; prox., proximal; n, number of cases; N, population size.
[1] Only first colonoscopy per screenee.
[2] Adequate resection means snare polypectomy of traditional serrated adenoma, hyperplastic polyp, sessile serrated polyp, non-advanced adenoma, and advanced adenoma > 3 mm.
[3] Moderate and severe pain defined as painful colonoscopy, 2013 excluded due to missing identification.
[4] In 2012 only possible to answer yes or no, therefor not included and 2013 not included due to missing participant identification.

▶ **Table 3** Comparison of colonoscopy key performance indicators between consultants and trainees.

| Quality indicators | Trainees | Consultants | *P* (univariate) | *P* (multivariate) |
|---|---|---|---|---|
| No. colonoscopies, N (row %) | 6,655 (87.8) | 921 (12.2) | – | – |
| No. of 1st colonoscopies per participant, N (row %) | 6,147 (88.8) | 778 (11.2) | – | – |
| Cecum intubated, n/N (%)[1] | 6,003/6,147 (97.7) | 749/778 (96.3) | 0.02 | 0.02 |
| Adenoma detected after FIT, n/N (%)[1] | 2,414/4,192 (57.6) | 253/503 (50.3) | <0.01 | <0.01 |
| Advanced adenoma detected after FIT, n/N (%)[1] | 1,042/4,192 (24.9) | 129/503 (25.7) | 0.70 | 0.07 |
| Proximal adenoma detected after sigmoidoscopy n/N (%)[1] | 804/1,955 (41.1) | 82/275 (29.8) | <0.01 | <0.01 |
| Proximal AA detected after sigmoidoscopy, n/N (%)[1] | 169/1,955 (8.6) | 23/275 (8.4) | 0.88 | 0.97 |
| Cecum intubation time, mean (95 % CI), min[2] | 13.7 (13.4–13.9) | 12.6 (12.0–13.2) | <0.01 | <0.01 |
| Withdrawal time, mean (95 % CI), min[3] | 12.8 (12.5–13.2) | 10.8 (9.6–11.9) | <0.01 | <0.01 |
| Adequate resection, n/N (%)[4] | 9,734/10,254 (94.9) | 1,122/1,205 (93.1) | 0.01 | 0.01 |
| Serious adverse events, n/N (%) | | | | |
| ▪ Perforation | 5/6,655 (0.08) | 1/921 (0.11) | 0.54 | NA |
| ▪ Bleeding | 38/6,655 (0.57) | 11/921 (1.19) | 0.03 | 0.05 |
| ▪ Death | 0/6,655 (0.00) | 1/921 (0.11) | 0.13 | NA |
| Total adverse events | 43/6,655 (0.65) | 13/921 (1.41) | 0.01 | 0.02 |
| Pain, moderate/severe, n/N (%)[5] | 924/4,353 (21.2) | 99/406 (24.4) | 0.14 | 0.16 |
| Severe pain, n/N (%)[5] | 337/4,353 (7.7) | 35/406 (8.6) | 0.53 | 0.55 |
| Very satisfied, n/N (%)[6] | 3,866/4,202 (92.0) | 324/353 (91.8) | 0.88 | 0.82 |

NA, not applicable; n, number of cases; N, population size.
[1] Only first colonoscopies per screenee.
[2] Only complete colonoscopies with registered time to reach cecum (n = 6,063 performed by trainees, n = 695 performed by consultants).
[3] Only complete diagnostic colonoscopies (n = 1,337 performed by trainees, n = 177 performed by consultants).
[4] Snare polypectomy of traditional serrated adenoma, hyperplastic polyp, sessile serrated polyp, non-advanced adenoma, and advanced adenoma > 3 mm.
[5] 2013 excluded due to missing participant identification.
[6] In 2012 only possible to answer yes or no, therefor not included and 2013 not included due to missing participant identification.

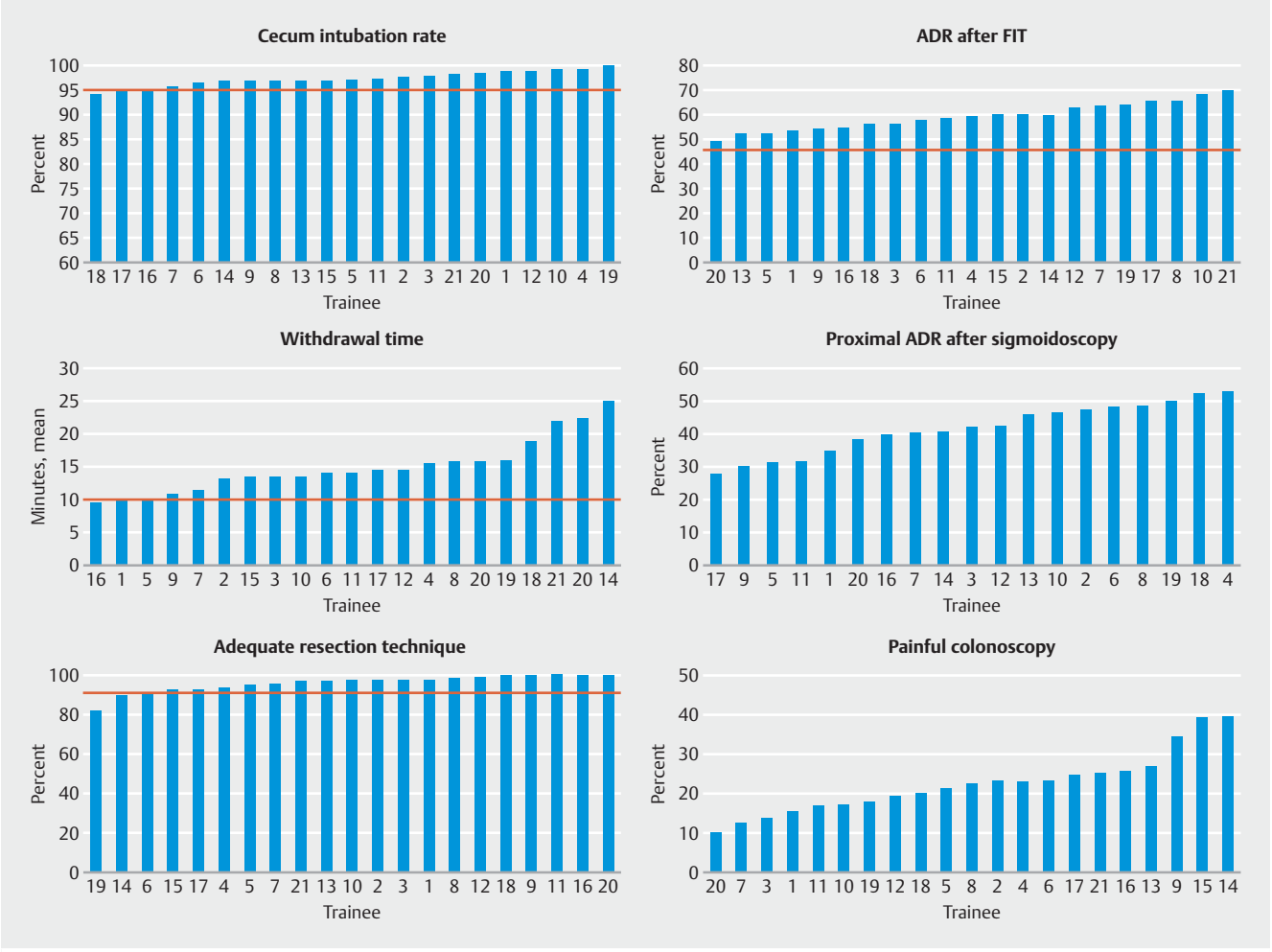
for proximal ADR ranging from 27.8 % to 52.5 % both at univariate analysis ($P$ < 0.01) and multivariable analysis ($P$ < 0.01). Proximal AADRs were borderline different between endoscopists at univariate analysis (1.3 %-24.0 %, $P$ = 0.09) and significantly different at multivariable analysis ($P$ = 0.03). The mean withdrawal time was 12.8 min (95 % CI 12.5min-13.2 min) (range 9.6 min-25.1 min) ($P$ < 0.01 at univariate and multivariable analysis). In total, 94.9 % of lesions larger than 3 mm were removed adequately by snare polypectomy but varied between 81.0 % and 100 % per trainee ($P$ < 0.01 at univariate and multivariable analysis). The proportion of painful colonoscopies was 21.2 % (range 9.8 %-39.2 %), with significant inter-endoscopist variation between trainees both at univariate and multivariable analysis ($P$ < 0.01), while patient satisfaction did not differ significantly.

## Learning curves for trainees as a group

Learning curves per year of screening after the training period showed that trainees as a group during the first three years after training shortened their cecum intubation time and withdrawal time, improved their adequate polyp resection rate, and reduced the rate of painful colonoscopies (all $P$ < 0.01) (▶ **Table 4**, **Supplementary Fig. 1**). Patient satisfaction improved ($P$ = 0.05). Over time, there was a negative trend for ADR in FIT screening for the trainees, but when only considering colonoscopies performed after a positive FIT in the first screening round, the ADR was unchanged. CIR and proximal ADR in sigmoidoscopy screening were unchanged over time.

## Comparison trainees to consultants as groups

After training, trainees performed 6.655 (87.8 %) screening colonoscopies, and consultants performed 921 colonoscopies (12.2 %) (▶ **Table 3**). The CIR of trainees was higher (97.7 % for trainees vs. 96.3 % for consultants, univariate $P$ = 0.02, multivariable $P$ = 0.02, OR 1.6, 95 % CI 1.1–2.5) (**Supplementary Table 1**). The trainees detected more adenomas in FIT screening (57.6 % vs. 50.3 %, univariate and multivariable $P$ < 0.01, OR 1.4, 95 % CI 1.2–1.8) and proximal adenomas after sigmoidoscopy screening (41.1 % vs. 29.8 %, univariate and multivariable $P$ < 0.01, OR 1.6, 95 % CI 1.2–2.1) than consultants. There was no difference between trainees and consultants in AADR after




▶ **Fig. 2** Comparison of colonoscopy key performance indicators among trainees. ADR, adenoma detection rate; FIT, fecal immunochemical test. Red horizontal lines illustrate international target standards (CIR: Target standard ≥ 95 % [1]. ADR FIT: recommended ADR > 45 % in men and > 35 % in women [22]. Adequate resection rate: target standard ≥ 90 % [1]. Withdrawal time: Target standard mean 10 min [1]).

a positive FIT screening and proximal AADR after a positive sigmoidoscopy. Both mean cecum intubation time (13.7 min vs 12.6 min, univariate and multivariable $P < 0.01$) and mean withdrawal time were longer (12.8 min vs. 10.8 min, univariate and multivariable $P < 0.01$) in trainees and adequate polyp resection rate was higher in trainees compared to consultants (94.9 % vs. 93.1 %, univariate and multivariable $P = 0.01$). Fewer SAEs occurred in trainees' colonoscopies than consultants' (0.65 % vs. 1.41 %, univariate $P = 0.01$, multivariable $P = 0.02$, OR 0.5, 95 % CI 0.3–0.9). Rates of painful colonoscopies and rates of very satisfied participants were similar (table 3, supplementary table 2). Neither was there a difference in the proportion of colonoscopies with severe pain between trainees and consultants.

## Discussion

This study is the first to show how important KPIs continuously improved after initial training and that high-quality colonoscopy performance reaching international target standards was achieved. Trainees also outperformed consultants on several KPIs. Nevertheless, there were significant inter-endoscopist variations for most of the KPIs. This study might be helpful when scheduling assessment of recently trained endoscopists.

In our study, trainees as a group achieved a higher CIR than consultants. All but two trainees, were above the recommended 95 % target for CIR in screening colonoscopies [1, 2] confirming that most trainees can perform complete colonoscopies independently after an intensive training consisting of 300 training colonoscopies [10]. In accordance, a previous study reported that intensive training was associated with achieving complete colonoscopy (reaching caecum, terminal ileum, ileo-colon anastomosis or neo terminal ileum) without assistance from the trainer [14]. Still, it is important to emphasize that continuous assessment of KPIs is required to make it possible to act on and assist underperforming endoscopists to improve.

All trainees reached ADRs in FIT-positive colonoscopies above the recommended target [22], although the highest-ranking trainee had a 1.4-fold higher ADR in FIT-positives and a 1.9-fold higher proximal ADR after a positive sigmoidoscopy

▶ **Table 4** Learning curves for key performance indicators among trainees.

| Quality indicator | First year | Second year | Third year | >Third year | *P*[1] |
|---|---|---|---|---|---|
| Cecum intubation rate, n/N (%)[2] | 2,482/2,541 (97.7) | 1,531/1,573 (97.3) | 1,107/1,133 (97.7) | 883/900 (98.1) | 0.21 |
| Proximal ADR after sigmoidoscopy, n/N (%)[2] | 339/825 (41.1) | 210/532 (39.5) | 157/375 (41.9) | 98/223 (44.0) | 0.87 |
| ADR after FIT,n/N (%)[2] | 1,025/1,716 (59.7) | 606/1,041 (58.2) | 430/758 (56.7) | 353/677 (52.1) | <0.01 |
| ADR after FIT 1. round, n/N (%)[2] | 522/858 (60.8) | 260/412 (63.1) | 101/163 (62.0) | 51/90 (56.7) | 0.33 |
| Mean cecum intubation time, min (95 %CI)[3] | 14.6 (14.2–15.0) | 14.0 (13.5–14.5) | 12.1 (11.6–12.6) | 12.5 (11.9–13.0) | <0.01 |
| Mean withdrawal time, min (95 % CI)[4] | 13.5 (12.9–14.2) | 13.5 (12.9–14.2) | 12.0 (11.4–12.6) | 11.3 (10.7–11.9) | <0.01 |
| Adequate resection technique, n/N (%)[5] | 4,030/4,319 (93.3) | 2,669/2,791 (95.6) | 1,865/1,926 (96.8) | 1,170/1,218 (96.1) | <0.01 |
| Very satisfied, n/N (%)[6] | 1,437/1,569 (91.6) | 988/1,083 (91.2) | 825/892 (92.5) | 616/658 (93.6) | 0.05 |
| Painful colonoscopy, n/N (%)[7] | 387/1,706 (22.7) | 235/1091 (21.5) | 188/897 (21.0) | 114/659 (17.3) | <0.01 |

ADR, adenoma detection rate; FIT, fecal immunochemical test; n, number of cases; N, population size
[1] Tests changes in time as continuous variable and adjusted for covariates as explained in the method section.
[2] Only first colonoscopies per screenee.
[3] Only complete colonoscopies with registered time to reach cecum (n = 6,063).
[4] Only complete diagnostic colonoscopies (n = 1,337).
[5] Snare polypectomy of traditional serrated adenoma, hyperplastic polyp, sessile serrated polyp, non-advanced adenoma, and advanced adenoma > 3 mm.
[6] In 2012 only possible to answer yes or no, therefor not included and 2013 not included due to missing participant identification.
[7] Moderate and severe pain defined as painful colonoscopy, 2013 excluded due to missing participant identification.

compared to the lowest-ranking trainee, in line with data from a screening sigmoidoscopy trial [23]. In accordance with a retrospective study [24], trainees detected more adenomas than consultants. Importantly, in our study there was no difference in the detection of advanced adenomas between trainees and consultants. This may indicate that trainees inspect the mucosa more thoroughly, possibly because specifically trained to perform screening colonoscopies in asymptomatic participants, in contrast to consultants in routine clinics primarily looking for changes explaining symptoms. To the best of our knowledge no previous study reports ADR changes over time. Decreased ADR in FIT screening over time, caused by decreased prevalence of adenomas in subsequent screening rounds, disappeared when adjusting for FIT rounds.

Trainees performed cecum intubation and withdrawal faster with increasing experience but still above ten minutes for withdrawal and with unchanged ADR and decreasing pain scores, in line with a previous study showing a limited increase in ADR for withdrawal time above 10 minutes [25]. Shortening withdrawal time and cecum intubation time improves cost-effectiveness and may be pursued if it has no negative consequences on other KPIs as pain and ADR.

European guidelines recommend that ≥ 90 % of lesions larger than 3 mm should be removed by snare polypectomy [1]. Recently, cold snare polypectomy was found to be safe, effective, and cheap for removal of lesions smaller than 10 mm [26], but not validated and implemented the first years of our trial, therefore the rate of cold snare polypectomy was not assessed. All but two trainees reached the target standard for snare polypectomy, but still the inter-endoscopist variation was significant. As far as we know, this study is the first to report changes in adequate polypectomy rate over time and to show an increasing proportion of adequately removed lesions with increasing experience.

Another important finding was the inter-endoscopist variation in the proportion of painful colonoscopies with up to a 4-fold difference in pain rate between trainees, as pain is a barrier to colonoscopy [27, 28]. However, it is not easy to set a benchmark for acceptable proportion of painful colonoscopies since the acceptance of discomfort and the expectations regarding sedation may depend heavily on cultural expectations. The number of painful colonoscopies decreased with increased experience and emphasizes the need to focus on painless colonoscopy technique avoiding looping and thus stretching of the mesentery. Use of water-assisted techniques and ultrathin colonoscopes are associated with less pain during colonoscopy [29, 30]. Unfortunately, we have not assessed these factors systematically throughout the trial.

Most participants were very satisfied with their colonoscopy experience, the inter-endoscopist variation in patient-reported satisfaction was low and there was no difference between trainees and consultants. Patient satisfaction improved over time with increasing experience of trainees. It is possible that satisfaction is a better marker than pain for the acceptance of a colonoscopy in patients accepting a certain degree of short-term pain. This may be particularly true since sedation is a known barrier to screening colonoscopies [31, 32].

SAEs rates were in line with those reported from the gFOBT-based English screening program [33]. Trainees caused less significant bleedings than consultants even when adjusting for

consultants possibly performing the most advanced polypectomies. The question of a possibly more liberal use of prophylactic clips among trainees remains unanswered since this was not systematically monitored. As SAEs were rare, we did not investigate time trends or SAE rates per individual trainee.

An important strength of this study is the high number of trainees and the long-term assessment of several KPIs while previous studies often focus only on one KPI. Furthermore, the screening database provided high-quality data. The availability of patient experience and outcome measures is an additional strength.

A limitation of the present study is missing details on colonoscopy training. Consultants at both endoscopy centers had attended courses based on JAG's TCT courses and upskill training in colonoscopy. The courses probably improve the trainer's pedagogic skills for supervision and performance enhancing feedback to trainees [34]. However, consultants not specifically trained in TCT courses also supervised the trainees, but as shown in a previous study, TCT courses have a significant effect on overall center performance. This may suggest that the colonoscopy training skills of all of a center's consultants are improved if at least one consultant attends TCT courses [35].

Unfortunately, we could not assess individual learning curves as the total numbers for KPIs for individual trainees were too small and therefore changes over time are only reported for trainees as a group. However, we observed significant inter-endoscopist variation in trainees and a few trainees did not reach the target standards for all KPIs. Therefore, trainee's colonoscopy performance should be monitored after initial training and individualized supervision and training should be provided if there is uncertainty about the achievement of adequate colonoscopy quality.

## Conclusions

Trainees continued to improve their colonoscopy performance for several KPIs after initial training, they performed high-quality procedures, and were superior to consultants for several KPIs, demonstrating that endoscopists can be successfully trained within an intensive training program. The improvement might be explained by a paradigm shift in endoscopists' training curriculum with high focus on conscious competence and competency-based training, continuous access to individual KPIs and biannual meetings comparing endoscopists' KPIs and providing individual feedback. The study also indicate that it is necessary to schedule assessments of the trainees even after initial training. Still, KPIs varied between trainees. Further studies are needed to assess root causes for the variation between trainees, varying learning curves, and why consultants perform less well than experienced trainees for several KPIs.

## Acknowledgments

The bowel preparation was provided free of charge by Ferring Pharmaceuticals; none of the authors disclose relationships or activities that could appear to have influenced the submitted work. This work was supported by the Norwegian Parliament (Norwegian national budget from 2011) in preparation for a future nationwide colorectal cancer screening program. The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The authors thank Roland Valori, MD, FRCP, Gloucestershire Royal Hospital, John Anderson, MD, FRCP, Cheltenham General Hospital, and Professor Donald MacIntosh, Dalhousie University Hospital, for their immense contribution to endoscopy training and creation of the endoscopy training system in Norway. They also thank the secretaries, nurses, and physicians at the two endoscopy units for their considerable contribution to the screening trial.

## Competing interests

The authors declare that they have no conflict of interest.

## Funding

Norwegian Parliament

## Clinical trial

ClinicalTrials.gov (http://www.clinicaltrials.gov/)
NCT01538550
**TRIAL REGISTRATION:** Cross-sectional study
NCT01538550 at ClinicalTrials.gov (http://www.clinicaltrials.gov/)

## References

[1] Kaminski MF, Thomas-Gibson S, Bugajski M et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) quality improvement initiative. United Europ Gastroenterol J 2017; 5: 309–3342

[2] Rex DK, Schoenfeld PS, Cohen J et al. Quality indicators for colonoscopy. Gastrointest Endosc 2015; 81: 31–53

[3] Rex DK, Bond JH, Winawer S et al. Quality in the technical performance of colonoscopy and the continuous quality improvement process for colonoscopy: recommendations of the U.S. Multi-Society Task Force on Colorectal Cancer. Am J Gastroenterol 2002; 97: 1296–1308

[4] Gavin DR, Valori RM, Anderson JT et al. The national colonoscopy audit: a nationwide assessment of the quality and safety of colonoscopy in the UK. Gut 2013; 62: 242–249

[5] 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

[6] Robertson DJ, Lieberman DA, Winawer SJ et al. Colorectal cancers soon after colonoscopy: a pooled multicohort analysis. Gut 2014; 63: 949–9567

[7] Tollivoro TA, Jensen CD, Marks AR et al. Index colonoscopy-related risk factors for postcolonoscopy colorectal cancers. Gastrointest Endosc 2019; 89: 168–176 e1638

[8] Shaukat A, Rector TS, Church TR et al. Longer withdrawal time is associated with a reduced incidence of interval cancer after screening colonoscopy. Gastroenterology 2015; 149: 952–957

[9] Ekkelenkamp VE, Dowler K, Valori RM et al. Patient comfort and quality in colonoscopy. World J Gastroenterol 2013; 19: 2355–2361

[10] Faulx AL, Lightdale JR. ASGE Standard of Practice Committee. et al. Guidelines for privileging, credentialing, and proctoring to perform GI endoscopy. Gastrointest Endosc 2017; 85: 273–281

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

[12] Patwardhan VR, Feuerstein JD, Sengupta N et al. Fellowship colonoscopy training and preparedness for independent gastroenterology practice. J Clin Gastroenterol 2016; 50: 45–51

[13] Sedlack RE, Coyle WJ. Group ACE Research Group. Assessment of competency in endoscopy: establishing and validating generalizable competency benchmarks for colonoscopy. Gastrointest Endosc 2016; 83: 516–523 e511

[14] Ward ST, Mohammed MA, Walt R et al. An analysis of the learning curve to achieve competency at colonoscopy using the JETS database. Gut 2014; 63: 1746–1754

[15] Rajendran A, Pannick S, Thomas-Gibson S et al. Systematic literature review of learning curves for colorectal polyp resection techniques in lower gastrointestinal endoscopy. Colorectal Dis 2020; 22: 1085–1100

[16] Nasca TJ, Philibert I, Brigham T et al. The next GME accreditation system–rationale and benefits. N Engl J Med 2012; 366: 1051–1056

[17] Ekkelenkamp VE, Koch AD, de Man RA et al. Training and competence assessment in GI endoscopy: a systematic review. Gut 2016; 65: 607–61518

[18] Randel KR, Schult AL, Botteri E et al. Colorectal cancer screening with repeated fecal immunochemical test versus sigmoidoscopy: baseline results from a randomized trial. Gastroenterology 2021; 160: 1085–1096 e1085

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

[20] Nagtegaal ID, Odze RD, Klimstra D et al. The 2019 WHO classification of tumours of the digestive system. Histopathology 2020; 76: 182–188

[21] Hoff G, Bretthauer M, Huppertz-Hauss G et al. The Norwegian Gastronet project: Continuous quality improvement of colonoscopy in 14 Norwegian centres. Scand J Gastroenterol 2006; 41: 481–487

[22] Robertson DJ, Lee JK, Boland CR et al. Recommendations on fecal immunochemical testing to screen for colorectal neoplasia: a consensus statement by the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology 2017; 152: 1217–1237 e1213

[23] Bretthauer M, Skovlund E, Grotmol T et al. Inter-endoscopist variation in polyp and neoplasia pick-up rates in flexible sigmoidoscopy screening for colorectal cancer. Scand J Gastroenterol 2003; 38: 1268–1274

[24] Gianotti RJ, Oza SS, Tapper EB et al. A longitudinal study of adenoma detection rate in gastroenterology fellowship training. Digest Dis Sci 2016; 61: 2831–2837

[25] Lee TJ, Blanks RG, Rees CJ et al. Longer mean colonoscopy withdrawal time is associated with increased adenoma detection: evidence from the Bowel Cancer Screening Programme in England. Endoscopy 2013; 45: 20–26

[26] Kawamura T, Takeuchi Y, Asai S et al. A comparison of the resection rate for cold and hot snare polypectomy for 4-9 mm colorectal polyps: a multicentre randomised controlled trial (CRESCENT study). Gut 2018; 67: 1950–1957

[27] McLachlan SA, Clements A, Austoker J. Patients' experiences and reported barriers to colonoscopy in the screening context – a systematic review of the literature. Patient Educ Couns 2012; 86: 137–146

[28] Pylvanainen K, Kairaluoma M, Mecklin JP. Compliance and satisfaction with long-term surveillance in Finnish HNPCC families. Fam Cancer 2006; 5: 175–178

[29] Cadoni S, Leung FW. Water-assisted colonoscopy. Curr Treatment Options Gastroenterol 2017; 15: 135–154

[30] Garborg KK, Loberg M, Matre J et al. Reduced pain during screening colonoscopy with an ultrathin colonoscope: a randomized controlled trial. Endoscopy 2012; 44: 740–746

[31] Jetelina KK, Yudkin JS, Miller S et al. Patient-reported barriers to completing a diagnostic colonoscopy following abnormal fecal immunochemical test among uninsured patients. J Gen Int Med 2019; 34: 1730–173632

[32] Muthukrishnan M, Arnold LD, James AS. Patients' self-reported barriers to colon cancer screening in federally qualified health center settings. Prev medicine Rep 2019; 15: 100896

[33] Rutter MD, Nickerson C, Rees CJ et al. Risk factors for adverse events related to polypectomy in the English Bowel Cancer Screening Programme. Endoscopy 2014; 46: 90–97

[34] Hoff G, Botteri E, Huppertz-Hauss G et al. The effect of train-the-colonoscopy-trainer course on colonoscopy quality indicators. Endoscopy 2021; 53: 1229–1234

[35] 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