Thieme

Expertise in colonoscopy intubation does not predict diagnostic accuracy: a simulation-based study

OPEN ACCESS

Authors

Andreas S. Vilmann[1], Christian Lachenmeier[1], Morten Bo S. Svendsen[1], Bo Soendergaard[1,2], Yoon S. Park[3], Lars Bo Svendsen[1,4], Lars Konge[1,5]

Institutions

1   Rigshospitalet – CAMES, Copenhagen, Denmark
2   University Hospital Hvidovre, Department of Gastroenterology and Gastrointestinal Surgery, Copenhagen, Denmark
3   Harvard Medical School, Department of Medical Education Massachusetts Hospital, USA
4   Rigshospitalet – Surgical Gastroenterology, Copenhagen, Denmark
5   University of Copenhagen and the Capital Region of Denmark – Centre for Clinical Education, Copenhagen, Denmark

Corresponding author
Andreas Slot Vilmann, Rigshospitalet – CAMES, Blegdamsvej 9, Copenhagen 2100, Denmark
andreas.vilmann@gmail.com

ABSTRACT

Background and study aims Studies have linked cecal intubation rate with adenoma detection rate; however, the direct association between technical performance during colonoscopy intubation and withdrawal has never been explored. Thus, it remains unclear whether gentle and efficient intubation predicts superior mucosal inspection. The aim of this study was to investigate the correlation between performance during intubation and withdrawal in a simulation-based setup.

Methods Twenty-four physicians with various experience in colonoscopy performed twice on the Endoscopy Training System (ETS). Intubation skills were evaluated by assessing tasks on the ETS related to intubation (scope manipulation and loop management) and use of a computerized assessment tool called the 3D-Colonoscopy Progression Score (3D-CoPS). Diagnostic accuracy was defined by the number of polyps found during the ETS task of mucosal inspection. Pearson's correlations were calculated to explore associations between intubation skill and diagnostic accuracy.

Results The correlation analysis between 3D-CoPS and number of polyps found during mucosal inspection revealed a weak and insignificant correlation (0.157, $P=0.3$). Likewise, an insignificant correlation was seen between ETS intubation and number of polyps found (0.149, $P=0.32$).

Conclusions We found no evidence to support that technical performance during intubation is correlated with mucosal inspection performance in a simulation-based setting.

Introduction

Colonoscopy is the gold standard for diagnostics of the rectum and colon. Patient safety, diagnostic precision, and criteria for a successful colonoscopy are operator depended [1–3]. Therefore, increasing emphasis is being placed on demonstrating quality and competence for an individual endoscopist during intubation and withdrawal.

Quality indicators for intubation skills include cecal intubation rate (CIR), subjective assessment tools such as the direct observational procedural skills (DOPS), and computerized assessment tools [4, 5]. CIR is a retrospective measure that requires a large number of performed procedures to be a valid quality measure and DOPS is time-consuming for assessors and prone to observer bias. The 3D-Colonoscopy Progression Score (3D-CoPS), however, is a newly developed computerized assessment tool able to assess intubation skills automatically

and in an unbiased fashion, which was developed to overcome the downsides of current tools available, making it ideal [6].

Assessing mucosal inspection is primarily based on adenoma detection rates (ADR) and time spent during withdrawal [4, 7]. However, calculating the ADR requires a large number of procedures, which makes it inappropriate for assessing individual cases and trainees with few procedures [8]. Furthermore, differences in ADR have been reported even though a minimum withdrawal time was assured [9, 10]. This indicates that time as an independent quality factor does not guarantee a high mucosal inspection. Hence, technical skills associated with a high mucosal inspection still remain indefinite.

The association between performance during intubation and withdrawal has never been explored. Thus, it remains unclear whether expertise in intubation can predict expertise in mucosal inspection, and therefore, be used as a proxy parameter for overall colonoscopy expertise. The aim of this study was to investigate the correlation between performance during intubation and during withdrawal in a simulation-based setup.

Patients and methods

Participants

Twenty-four physicians with various experience were included. Participant demographics were: mean age 43 years (range 25–66), mean number of total colonoscopies performed 3286 (range 0–10,000), mean number of colonoscopies per year 228 (range 0–700) and mean years since graduation 15 (range 0–39) (▶ Table 1).

Equipment

All tests were performed using an Olympus colonoscope (CF-H180DL, Evis Exera II video center CV-180, Olympus Medical System Ltd, Tokyo, Japan), the Magnetic Endoscopy Imaging (MEI) system called the ScopeGuide (UPD-3, Olympus, Tokyo, Japan), and the Endoscopy Training System (ETS) (▶ Fig. 1) [11]. The ETS consists of two different physical simulators on which five different technical tasks are available. The first simulator is a linear model of the rectum and the sigmoid colon and is equipped with a simulated endoscopic tool, and audio and visual feedback. The first three tasks are performed on this model. The last two tasks are performed on the second simulator, which is a modified version of the Kyoto's Colonoscopy Training Model. The Colonoscopy Training Model is a realistic rubber model of the human colon with 10 polyps located throughout the rubber colon. The rubber colon can be modified into different forms and case 3 was chosen (Alpha loop in the sigmoid colon). Each participant performed twice on both simulators.

The first task, scope manipulation, reflects use of tip deflection and torque of the colonoscope. The goal is to align 10 numbered white triangles located within the model with two black triangles located externally on the video display of lumen. The second task, loop management, is navigation of the colonoscope from anus to cecum, during which the participant will encounter an alpha loop in the sigmoid colon. The purpose of this task is to reach the cecum within a time limit and to cor-

▶ Table 1 Demographics.

Participant number	Sex (female/male)	Age	Total colonoscopies	Colonoscopies the last year
1	M	65	5000	400
2	F	32	1	1
3	F	32	0	0
4	F	29	0	0
5	M	25	0	0
6	M	29	34	34
7	F	28	15	15
8	F	30	0	0
9	F	56	3000	300
10	M	60	10000	300
11	M	61	7500	550
12	M	32	5	5
13	M	31	4	4
14	F	28	0	0
15	F	30	30	30
16	M	53	10000	600
17	M	47	2000	700
18	M	41	4000	400
19	M	63	7000	1000
20	M	66	7000	500
21	M	48	7000	500
22	M	48	10000	400
23	F	28	0	0
24	M	59	8000	500

rectly manage the formation of a standard alpha loop. The task is scored by time and unsuccessful loop reduction leads to penalties. The third task, mucosal inspection, is inspection of the mucosa during withdrawal from the cecum to the anus. Ten identical polyps are located throughout the colon and as many as possible are to be found. The participants were not informed of the total number of polyps located inside the rubber colon. The three tasks were recorded by an investigator serving as the test proctor. If the time to complete each task surpassed the predetermined cutoff time, a timing score of zero was given. The key metrics and performance time then were used for calculation. Two tasks, tool targeting and retroflection, were left out since they do not represent technical aspects of intubation or withdrawal.
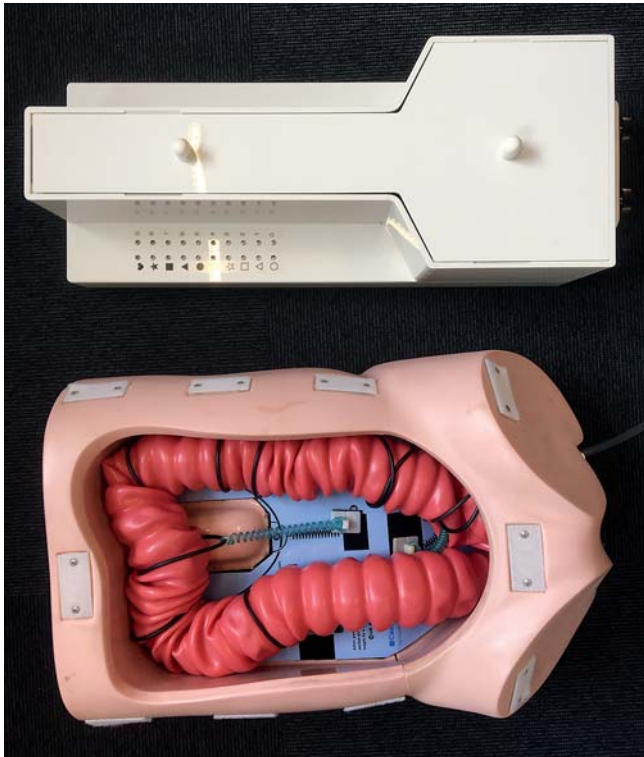
▶ **Fig. 1** The endoscopy training system (ETS).

3D-colonoscopy progression score (3D-CoPS):

The ScopeGuide system provides a real-time 3D image of the shape of the colonoscope inside the abdomen. The 3D image is based on XYZ-coordinates from electromagnetic coils placed along the length of the colonoscope. During task four, loop management, we collected XYZ-coordinates directly from the ScopeGuide system. Five different technical measures were applied to the XYZ-coordinates. Travel length is the distance of the tip traveled from anus to cecum. Tip progression is a calculation of how closely the tip of the colonoscope travel to the optimal pathway. Chase efficiency evaluates how closely a coil follows the same path as the coil in front. Shaft movement without tip progression refers to the situation where the tip of the colonoscope is lodged but the shaft is moving and potentially stretching the colon wall. Looping is quantification of how looped the colonoscope is during the procedure. The five measurements were built into a composite score, the 3D-CoPS, which is a computerized assessment tool that was developed to assess the technical performance of the progression from anus to cecum [6].

Outcome measures

The primary outcome was the correlation between intubation skills defined as 3D-CoPS and ETS tasks related to intubation (scope manipulation and loop management, referred to as ETS-intubation), and the number of polyps found during task 5. Secondary outcomes were the correlation between total colonoscopy experience, colonoscopies per year, time spent during mucosal inspection and number of polyps found during task 5.

Ethics

The regional committee of ethics evaluated and approved the study (H-17040471). All participants were provided with oral and written information regarding the trial. Participation was voluntary; no material goods were donated to the participants. The trial was registered (December 22, 2017) at clinicaltrials.gov with trial identification number NCT03401723.

Statistics

Statistical analysis was done with IBM SPSS statistics (PASW, version 22; SPSS Inc, Chicago, Ill). The level of statistical significance was set at $\alpha < .05$ for all tests. Bivariate correlations were used to test the expectation that intubation performance based on 3D-CoPS and ETS-intubation would be associated with one another and correlated with number of polyps found during the task of mucosal Inspection. Reliability was explored with the test-retest method [12].

Results

The study was conducted from October 2017 to February 2018 and included 24 participants. One participant was excluded due to unintended movement of the phantom model during the task loop management, which affected the data gathered from the ScopeGuide system, and therefore, the 3D-CoPS. A total of 46 completed data sets were used for the analysis (▶ **Table 2**).

A Pearson correlation was done, and a strong test-retest reliability was revealed between 3D-CoPS ($r = 0.86$, $P < 0.001$), ETS-intubation ($r = 0.83$, $P < 0.001$) and time spent during withdrawal ($r = 0.71$. $P < 0.001$). Number of polyps found had a test-retest reliability of 0.54 ($P = 0.008$). As expected, 3D-CoPS and ETS-intubation skills had a moderate to strong correlation (0.58, $P < 0.001$) (▶ **Fig. 2**).

Primary outcomes

The correlation analysis between 3D-CoPS and number of polyps found revealed a weak, insignificant correlation (0.16, $P = 0.297$) (▶ **Fig. 3**). Likewise, a weak correlation was seen between ETS-intubation, and number of polyps found (0.15, $P = 0.322$) (▶ **Fig. 4**).

Secondary outcomes

No correlation was found between total colonoscopy experience and time spent during withdrawal or number of polyps found. The correlation between time spent during withdrawal and colonoscopies per year was weak to moderate, meaning that increasing experience meant decreasing time spent during mucosal inspection (▶ **Table 3**).

Discussion

Studies have linked cecal intubation rate with ADR; however, the direct association between technical performance during colonoscopy intubation and withdrawal has never been explored. Thus, it remains unclear whether intubation skills pre-

► Table 2 Results of the 46 completed data sets.

Partici-pant number	Tip pro-gression (mean)	Travel length (mean)	Chase effi-ciency (mean)	Shaft move-ment without tip progres-sion (mean)	Looping (mean)	3D-CoPS (mean)	Number of polyps found (mean)	Time spent inspecting mucosa (mean)	ETS-in-tuba-tion (mean)
1	13045	4258	161	.81	457	−.24	8	120	219
2	66209	10548	232	.54	249	.41	7.5	433	94
3	60401	13483	248	.56	438	.67	7.5	428	126
4	49003	10910	210	.66	1620	.91	5	226	−6
5	6753	3949	211	.75	477	−.24	7	67	79
6	8189	2610	236	.80	238	−.25	9	67	228
7	26117	6341	210	.51	1206	.03	5.5	257	51
8	28160	5171	257	.69	120	-.04	8.5	261	83
9	9960	4936	174	.80	288	−.26	8	83	239
10	8881	2790	204	.86	154	−.28	6	75	173
11	33393	6812	159	.54	321	−.34	5.5	196	169
12	37438	6550	222	.53	262	−.16	5.5	277	185
13	39138	12486	414	.59	97	.81	7.5	286	186
14	45024	7547	352	.67	1488	1.08	7	333	107
15	10063	2561	240	.80	610	−.07	7	87	172
16	9848	3722	215	.85	311	−.12	9.5	94	215
17	10878	3093	292	.67	263	−.20	5.5	63	250
18	7741	2273	157	.76	71	−.63	5.5	45	236
20	23591	2854	139	.78	101	−.45	6.5	108	213
21	24177	4536	158	.83	263	.40	5	94	236
22	11592	2818	226	.90	159	.16	7.5	74	168
23	33682	6231	166	.71	593	.02	6	414	92
24	14588	3527	179	.79	88	−.38	7.5	95	229

Participant number 19 was excluded due to unintended movement of the phantom model during the task of loop management, which affected the data gathered from the ScopeGuide system, and therefore, the 3D-CoPS.
3D-CoPS; 3D-colonoscopy progression score; ETS-intubation, endoscopy training system-intubation.

dict proficiency in the skills related to a better mucosal inspection, i. e. higher ADR.

We used two different assessment tools for simulated colonoscopy (ETS-intubation and 3D-CoPS) to investigate the correlation between performance skills related to intubation and diagnostic accuracy based on number of polyps found. We found no correlation between intubation-related performance (based on 3D-CoPS and ETS-intubation) and number of polyps found during the task mucosal inspection. Nor was clinical experience associated with a higher polyp detection.

Our findings are in line with previous studies investigating transfer of skills from one procedure to another [13–15]. These studies investigated two different procedures, but the colonoscopy procedure should also be considered as two separate procedures with two different technical approaches needed to

succeed. Transfer of skills from simulation-based training to the clinic has shown improved outcomes in the early stages of training in the clinic [16], but no one has investigated transfer of mucosal inspection techniques in a simulation-based setting and training the skills needed remain unclear [17].

In 2013 Ritter et al. published a study aimed at gathering evidence of validity of a non-virtual-reality assessment tool, which later was added as a new task and today is called ETS [11, 18]. Results from these studies indicate that the experienced participants in our study had a lower score for mucosal inspection and found fewer polyps during withdrawal. A slightly different setup in the prior studies may be the reason for the differences. In task 5 (mucosal inspection) Ritter et al. describe the use of 10 to 20 polyps located throughout the rubber model, but in our study, the polyps were fixed to 10 polyps and loca-
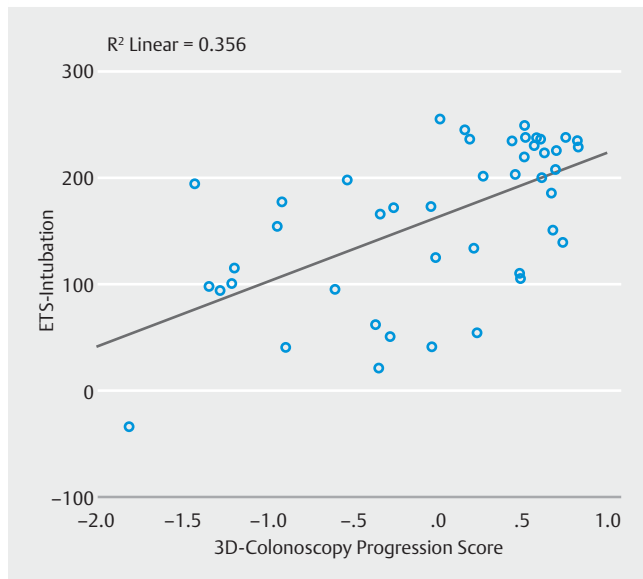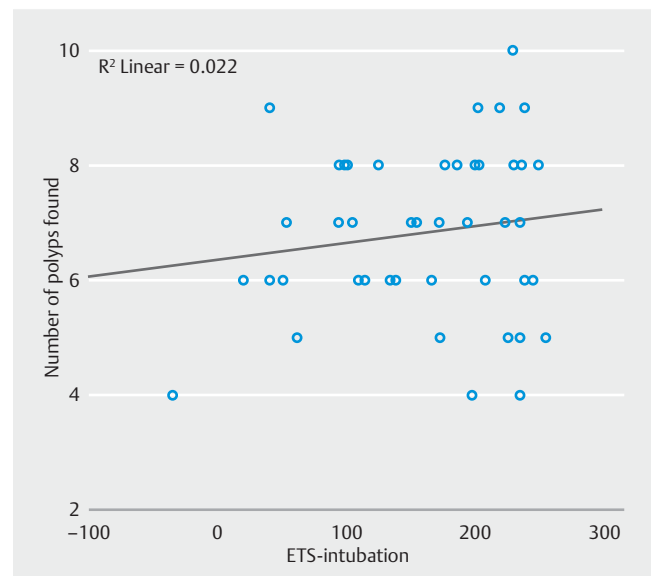
▶ Fig. 2 Correlation between 3D-CoPs and ETS-intubation (scope manipulation and loop management). 3D-CoPS, 3D-Colonoscopy Progression Score. ETS, Endoscopy Training system.
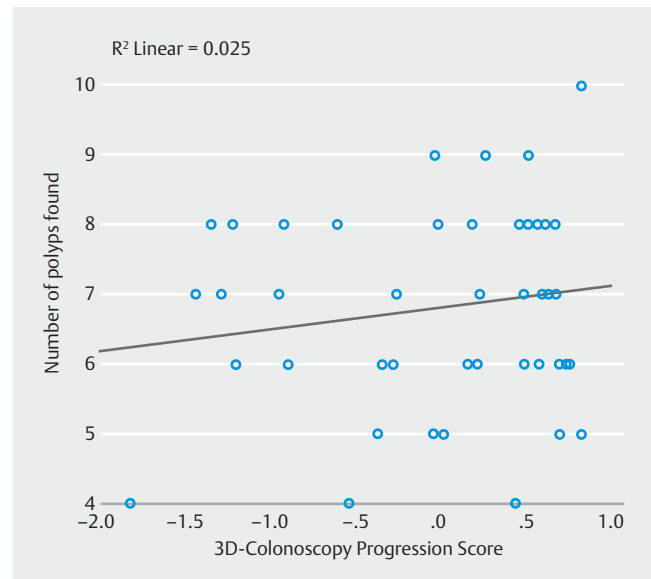
▶ Fig. 3 Correlation between 3D-CoPS and number of polyps found during ETS task Mucosal Inspection. 3D-CoPS, 3D-Colonoscopy Progression Score. ETS, Endoscopy Training System.

tion. Furthermore, in relation to aspects of polyp detection, the experience was not associated with finding more polyps. Real polyp recognition skills have revealed experienced-novice differences [19]. However, in the present study, all 10 polyps were identical and only the search component of polyp detection was tested. Hence, clinical experience with subtle features of polyps presented no specific advantage.

Mucosal inspection time had a weak negative but significant correlation with colonoscopies per year and ETS-intubation score, meaning that participants who performed well on the ETS-intubation tasks spent less time on mucosal inspection. This is in line with a previous study investigating mucosal inspection technique in a simulation-based setting [20]. They found that clinically experienced endoscopists tended to visualize more of the mucosa and at a faster rate. In a clinical study investigating withdrawal technique among 11 endoscopists with varying ADR, they found no difference in time spent during withdrawal but higher withdrawal technique scores among endoscopists with a moderate to high ADR. They concluded that withdrawal techniques may be more important compared with withdrawal time in predicting ADRs [9]. Likewise, a multicenter randomized trial explored the effect of an improvement program on ADR and found increasing ADR but unchanged withdrawal time, indicating that a good inspection technique is imperative for a high ADR [10]. The difference in mucosal time spent during withdrawal could be explained by the results of a prospective study demonstrating that experts with the lowest ADR tended to have a slow withdrawal speed but failed to adequately examine the back of the colonic folds for potentially missed polyps. In contrast, an endoscopist with a high ADR was more apt to carefully "reinsert" the scope in missed areas and might explain the difference in time spent during withdrawal [21]. Most participants in our study spent less time

▶ Fig. 4 Correlation between ETS-intubation (scope manipulation and loop management) and number of polyps found during the ETS task of mucosal inspection. ETS, Endoscopy Training System.

than the recommended minimum withdrawal time in the clinic for patients without polyps (mean mucosal inspection time; 388 seconds (SD 126), clinical minimal recommendation; < 360 to 600 seconds) [4, 7]. Reasons could be that some experts did not take the task of finding simulated polyps seriously enough due to lack of realism and consequences of missing a polyp. In two studies exploring performance in simulated mastoidectomy on a cadaveric and a virtual reality (VR) simulator, results demonstrated inappropriate behavior among experts concluding that simulator fidelity is a contributing factor [22].

Table 3 Bivariate correlation between 3D-CoPS, ETS-intubation (scope manipulation and loop management), number of polyps found, time spent inspecting mucosa, colonoscopies per year and total colonoscopy experience.

		3D-CoPS	ETS-intubation	Number of polyps found	Time spent inspecting mucosa	Colonoscopies per year	Total colonoscopy experience
3D-CoPS	Pearson Correlation	1	.596[1]	.157	−.246	.581[1]	.615[1]
	Sig. (2-tailed)		.000	.297	.100	.000	.000
	N	46	46	46	46	46	46
ETS-intubation	Pearson Correlation		1	.149	−.382[1]	.659[1]	.493[1]
	Sig. (2-tailed)			.322	.009	.000	.001
	N		46	46	46	46	46
Number of polyps found	Pearson Correlation			1	.207	−.084	.051
	Sig. (2-tailed)				.168	.581	.736
	N			46	46	46	46
Time spent inspecting mucosa	Pearson Correlation				1	−.302[2]	−.185
	Sig. (2-tailed)					.042	.220
	N				46	46	46
Colonoscopies per year	Pearson Correlation					1	.804[1]
	Sig. (2-tailed)						.000
	N					46	46
Total colonoscopy experience	Pearson Correlation						1
	Sig. (2-tailed)						
	N						46

3D-CoPS, 3D-Colonoscopy Progression Score; ETS, Endoscopy Training system.
[1] Correlation is significant at the 0.01 level (2-tailed).
[2] Correlation is significant at the 0.05 level (2-tailed).

Strengths and limitations

It is important to acknowledge that results from a simulated environment can never be completely realistic. Endoscopy training simulators whether being physical or virtual all lack realism [23]. The Colonoscopy Training Model is a rubber model of the colon and for use, needs to be lubricated to decrease friction between the colonoscope and the rubber colon. Incorrectly applied lubrication to the rubber colon makes handling loops and passing flexures difficult due to friction. However, we used the ETS simulators with established evidence of validity and the standardized setup made comparisons easily feasible. When investigating correlation, the strength lies in the sample being heterogeneous, which was the case for the 23 participants (total colonoscopy experience: 0–10,000). We suspect the simplicity of the simulator (task 5, mucosal inspection) to be the reason for lack of correlation between technical intubation performance and mucosal inspection. It is well known that polyps may appear or disappear during withdrawal due to the high-rise folds and peristalsis. However, in our simulated setup, the rubber colon did not have peristaltic movements, hence providing the participant with a better view of the mucosa, which may lower the time-dependent effect during withdrawal. ETS tests multiple tasks and the simulator is a strong training and assessment tool, but the task of inspecting the mucosa needs further development to be of valid use. Polyp size, location, and morphology need more diversity for the simulator to differentiate between experts and novices in the skills of mucosal inspection. Moreover, studies have shown that assessing in multiple ways increases reliability and lowers bias, hence a version that includes different tasks or simulators may prove better for that purpose [24].

Currently, no transferability of mucosal inspection in simulation-based training has been explored. Testing transferability from simulation to the clinic poses a challenge because the current primary quality indicator is ADR, which requires a large number of procedures to be valid. Eliminating ADR as a quality factor requires an assessment tool closely correlated to ADR. A clinical study investigating the relationship between ADR and an assessment tool able to assess both in the simulation-based setup and the clinic, such as 3D-CoPS, may be the answer for predicting diagnostic accuracy.

Conclusions

In conclusion, we found no evidence to support that technical performance during intubation is correlated with mucosal inspection performance in a simulation-based setting. Clinical experience was correlated with less time spent on mucosal inspection, indicating higher efficiency in polyp detection but not in the number of polyps found. Our results indicate that more realistic simulators are needed for training and assessing competence in polyp detection.

Competing interests

The authors declare that they have no conflict of interest.

References

[1] Rabeneck L, Paszat LF, Hilsden RJ et al. Bleeding and perforation after outpatient colonoscopy and their risk factors in usual clinical practice. Gastroenterology 2008; 135: 1899–1906.e1

[2] Sedlack RE, Coyle WJ. Assessment of competency in endoscopy: Establishing and validating generalizable competency benchmarks for colonoscopy. Gastrointest Endosc 2016; 83: 516–523.e1

[3] Dafnis G, Granath F, Påhlman L et al. The impact of endoscopists' experience and learning curves and interendoscopist variation on colonoscopy completion rates. Endoscopy 2001; 33: 511–517

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

[5] Barton JR, Corbett S, Van Der Vleuten CP. The validity and reliability of a Direct Observation of Procedural Skills assessment tool: Assessing colonoscopic skills of senior endoscopists. Gastrointest Endosc 2012; 75: 591–597

[6] Vilmann AS, Lachenmeier C, Svendsen MBS et al. Using computerized assessment in simulated colonoscopy: a validation study. Endosc Int Open 2020; 08: E783–E791

[7] 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. Endoscopy 2015; 49: 1175–1189

[8] Do A, Weinberg J, Kakkar A et al. Reliability of adenoma detection rate is based on procedural volume. Gastrointest Endosc 2013; 77: 376–380

[9] Lee RH, Tang RS, Muthusamy VR et al. Quality of colonoscopy withdrawal technique and variability in adenoma detection rates (with videos). Gastrointest Endosc 2011; 74: 128–134

[10] 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). Gastrointest Endosc 2017; 85: 538–545.e4

[11] Ritter EM, Taylor ZA, Wolf KR et al. Simulation-based mastery learning for endoscopy using the endoscopy training system: a strategy to improve endoscopic skills and prepare for the fundamentals of endoscopic surgery (FES) manual skills exam. Surg Endosc Other Interv Tech 2018; 32: 413–420

[12] Downing S, Yudkowsky R. Assessment in Health Professions Education. Routledge; United Kingdom: 2009

[13] Sofia A, Thomsen S, Konge L et al. Is there inter-procedural transfer of skills in intraocular surgery ? A randomized controlled trial Acta Ophthalmol 2017; 57: 845–851

[14] Bjerrum F, Sorensen JL, Konge L et al. Randomized trial to examine procedure-to-procedure transfer in laparoscopic simulator training. Br J Surg 2016; 103: 44–50

[15] Mills J, Hougen H et al. Does robotic surgical simulator performance correlate with surgical skill? J Surg Educ 2017; 74: 1052–1056 doi:10.1016/j.jsurg.2017.05.011 (Epub 2017 Jun 13)

[16] Brydges R, Hatala R, Zendejas B et al. Linking simulation-based educational assessments and patient-related outcomes. Acad Med 2015; 90: 246–256

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

[18] Ritter EM, Cox TC, Trinca KD et al. Simulated Colonoscopy Objective Performance Evaluation (SCOPE): A non-computer-based tool for assessment of endoscopic skills. Surg Endosc Other Interv Tech 2013; 27: 4073–4080

[19] Hill A, Horswill MS, Plooy AM et al. Assessment of colorectal polyp recognition skill: development and validation of an objective test. Surg Endosc 2017; 31: 2426–2436

[20] Zupanc CM, Wallis GM, Hill A et al. Assessing colonoscopic inspection skill using a virtual withdrawal simulation: a preliminary validation of performance metrics. BMC Med Educ 2017; 17: 118

[21] Rex DK. Colonoscopic withdrawal technique is associated with adenoma miss rates. Gastrointest Endosc 2000; 51: 33–36

[22] Andersen SAW, Mikkelsen PT, Sørensen MS. Expert sampling of VR simulator metrics for automated assessment of mastoidectomy performance. Laryngoscope 2019: doi:10.1002/lary.27798

[23] Hill A, Horswill MS, Plooy AM et al. Assessing the realism of colonoscopy simulation: The development of an instrument and systematic comparison of 4 simulators. Gastrointest Endosc 2012; 75: 631–640. e3

[24] Konge L, Svendsen MBS, Preisler L et al. Combining different methods improves assessment of competence in colonoscopy. Scand J Gastroenterol 2017; 52: 601–605