

Using computerized assessment in simulated colonoscopy: a validation study



Authors

Andreas Slot Vilmann¹, Christian Lachenmeier¹, Morten Bo Søndergaard Svendsen¹, Bo Søndergaard², Yoon Soo Park³, Lars Bo Svendsen⁴, Lars Konge¹

Institutions

- 1 Rigshospitalet – CAMES, Copenhagen, Denmark
- 2 University Hospital Hvidovre, Department of Gastroenterology and Gastrointestinal Surgery, Copenhagen, Denmark
- 3 University of Illinois at Chicago, Department of Medical Education, Chicago, Illinois, United States
- 4 Rigshospitalet – Surgical Gastroenterology, Copenhagen, Denmark

submitted 29.4.2019

accepted after revision 27.1.2020

Bibliography

DOI <https://doi.org/10.1055/a-1132-5259> |
Endoscopy International Open 2020; 08: E783–E791
© Georg Thieme Verlag KG Stuttgart · New York
eISSN 2196-9736

Corresponding author

Andreas Slot Vilmann, Blegdamsvej 9, Copenhagen 2100, Denmark
Phone: +004538664631
andreas.vilmann@gmail.com

ABSTRACT

Background and study aims Patient safety during a colonoscopy highly depends on endoscopist competence. Endoscopic societies have been calling for an objective and regular assessment of the endoscopists, but existing as-

essment tools are time-consuming and prone to bias. We aimed to develop and gather evidence of validity for a computerized assessment tool delivering automatic and unbiased assessment of colonoscopy based on 3 dimensional coordinates from the colonoscope.

Methods Twenty-four participants were recruited and divided into two groups based on experience: 12 experienced and 12 novices. Participants performed twice on a physical phantom model with a standardized alpha loop in the sigmoid colon. Data was gathered directly from the Olympus ScopeGuide system providing XYZ-coordinates along the length of the colonoscope. Five different motor skill measures were developed based on the data, named: Travel Length, Tip Progression, Chase Efficiency, Shaft movement without tip progression, and Looping.

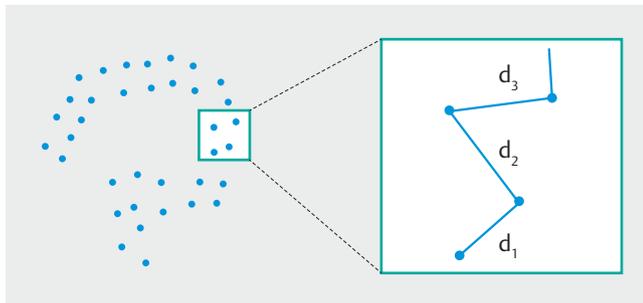
Results The experienced had a lower travel length ($P < 0.001$), tip progression ($P < 0.001$), chase efficiency ($P = 0.001$) and looping ($P = 0.006$), and a higher shaft movement without tip progression ($P < 0.001$) reaching the cecum compared with the novices. A composite score was developed based on the five measurements to create a combined score of progression, the 3D-Colonoscopy-Progression-Score (3D-CoPS). The 3D-CoPS revealed a significant difference between groups (experienced: 0.495 (SD 0.303) and novices -0.454 (SD 0.707), $P < 0.001$).

Conclusion This study presents a novel, real-time computerized assessment tool for colonoscopy, and strong evidence of validity was gathered in a simulation-based setting. The system shows promising opportunities for automatic, unbiased and continuous assessment of colonoscopy performance.

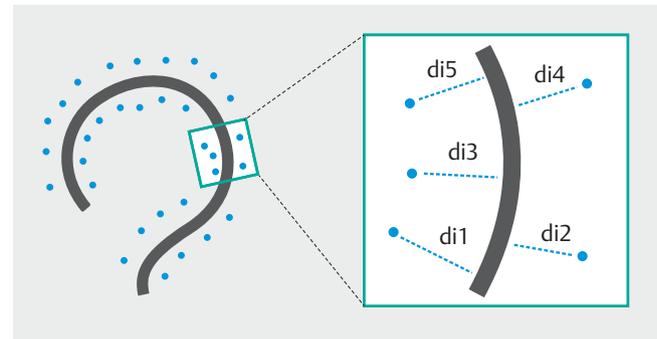
Introduction

Colonoscopy is a common procedure and considered the gold standard for the investigation of the colon and rectum. Cancer screening has led to an increase in colonoscopy procedures and thus a need for educating more endoscopists [1,2]. Learning the skills needed to perform a safe colonoscopy is not an easy task, and it is imperative to ensure that every single endos-

copist is competent before being allowed to practice independently [3]. The importance of regular assessment of individual endoscopists has been recognized by the major gastroenterology societies [4,5]. Recommended quality indicators are cecal intubation rate (CIR) above 90%, cecal intubation time (CIT) less than 16 minutes, and a minimum number of performed colonoscopies [2,6–8]. Unfortunately, many performed procedures are needed to make CIR and CIT statistically valid indicators for technical competence making them unsuitable for en-



► **Fig. 1** Travel length is measured by the cumulative distance in cm between all the data points from the first coil, i. e., the tip of the colonoscope.



► **Fig. 2** Tip progression is the cumulative distance in cm from all the data points from the first coil, i. e. the tip of the colonoscope, to the optimal path line defined by the shape of the colonoscope when the tip is positioned in the cecum.

asuring the competence of new endoscopists [9]. Furthermore, a specific number of performed procedures does not ensure that each endoscopist is competent [9]. To deliver high-quality care for patients, competence needs to be based on observation such as Gastrointestinal Endoscopy Competence Assessment Tool (GiECAT), Global Assessment of Gastrointestinal Endoscopy Skills (GAGES), Assessment of Competency in Endoscopy (ACE), or Direct Observation of Procedural Skills (DOPS) [10–14]. These assessment tools focus on technical performance and cognitive skills. Unfortunately, these are prone to observer bias and present a formidable workload for expert supervisors/assessors, in turn reducing compliance [2, 4, 15, 16]. Nerup et al. developed a computerized assessment tool, the Colonoscopy Progression Score (CoPS) to allow an automatic and unbiased assessment of colonoscopy skills [17]. The CoPS metric relied on computer-based image analysis of the ScopeGuide video and could only assess the progression of the colonoscope on its route from anus to cecum based on the two-dimensional positions of the tip. Direct access to XYZ-coordinates from the whole length of the colonoscope could allow the development of a system that assesses more than just the route of the tip, such as looping, tip control, and optimal path to ensure a safe and smooth progression. However, solid validity evidence proving meaningful measurements is necessary before the system can be trusted to assess competence in colonoscopy.

This study aimed to develop and collect validity evidence for a computerized assessment tool, called the 3D-Colonoscopy Progression Score (3D-CoPS).

Methods

Development of the assessment system

The system utilizes a Magnetic Endoscopy Imaging system called ScopeGuide (UPD-3, Olympus, Tokyo, Japan) where electromagnetic coils are built in along the length of the colonoscope. Each coil generates a magnetic field that is picked up by receiver coils in a receiver dish and processed in the UPD-3 device. For normal use, the UPD-3 system renders 2D images of the shape of the colonoscope to assist the operator in understanding and handling certain issues during a procedure. We collected XYZ-coordinates directly from the UPD-3 unit

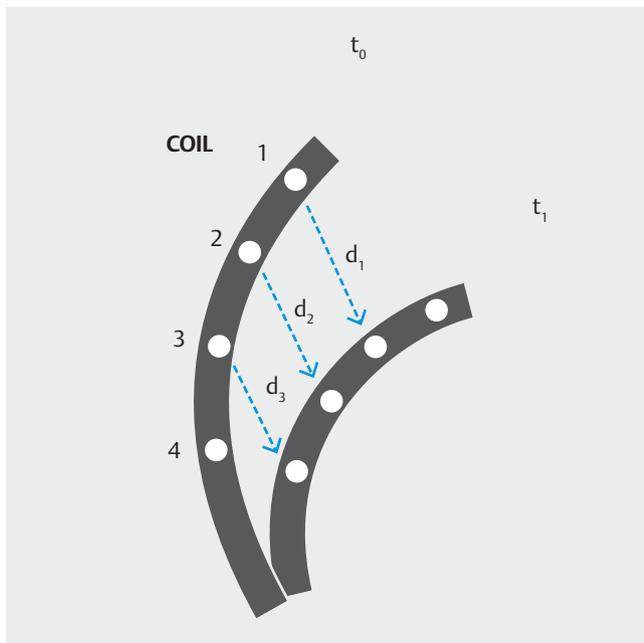
through an Olympus receiver box (UCES-3) five times per second (5 Hz) during insertion of the colonoscope.

The aim was to use the XYZ data to create a combined score of progression (3D-CoPS) based on five different measures: 1. Travel length; 2. Tip progression; 3. Chase efficiency; 4. Shaft movement without tip progression; and 5. Looping.

Travel length: The optimal path from anus to cecum reflects the shortest possible distance. The measurement is the travel length of the tip from anus to cecum provided in cm (► **Fig. 1**). Higher score values indicate a longer distance traveled whereas a low score represents a shorter path to reach cecum.

Tip progression: A score of progression has previously been described in a 2D format (Colonoscopy Progression Score) using coordinates of the tip's position during insertion [17]. This was based on tracking the video feed to get X, Y coordinates (e. g. the frontal plane). Getting data directly from the ScopeGuide system increases the precision because adding a third dimension (Z) provides information regarding the progression in the transverse plane of the phantom model. The Tip progression score is based on the assumption that the optimal path line is defined by the shape of the colonoscope when the tip is positioned in the cecum. The score is completed by calculating the minimal distance from all data points collected from tip to the optimal path line (► **Fig. 2**). A low score indicates a more smooth and even progress during insertion.

Chase efficiency: A straight colonoscope is in principle an instrument that will respond more instantly to both torque and in/out movements. This situation is ideal since the applied force needed to advance is less compared to a more bent colonoscope. When the shaft of the colonoscope starts to bend it is less likely to travel in the same path as the tip and the colonoscope is displaced. The measurement is the distance over time (meter per second) between neighbor coils meaning that if coil 1 travels from A to B and coil 2 travels from C to D then the calculated distance are the distance between B and D. The calculation is done for all the coils. The measurement evaluates how closely a coil follows the same path as the coil in front. The displacement is small (low score) if the coils follow one another closely and a large displacement (high score) represents a

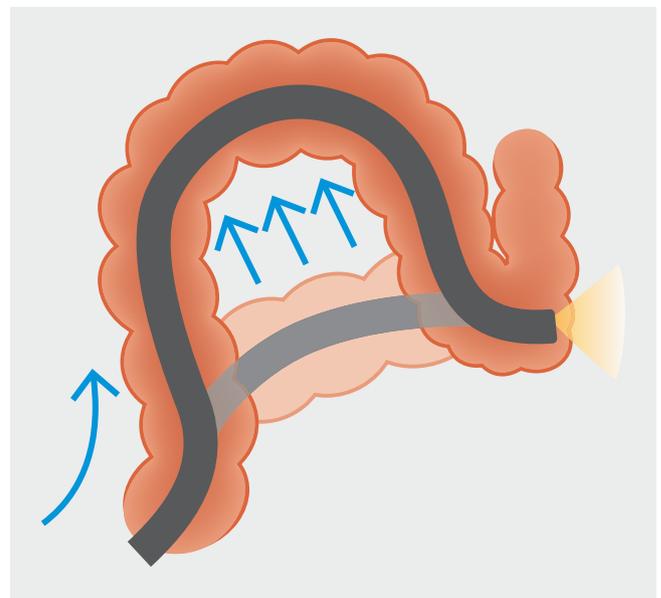


► **Fig. 3** Chase efficiency is a score of how closely a coil follows the same path as the coil in front (meter per second). A shift away from the coil in front increases the score and represents a greater distance between the coils from one data point to the next.

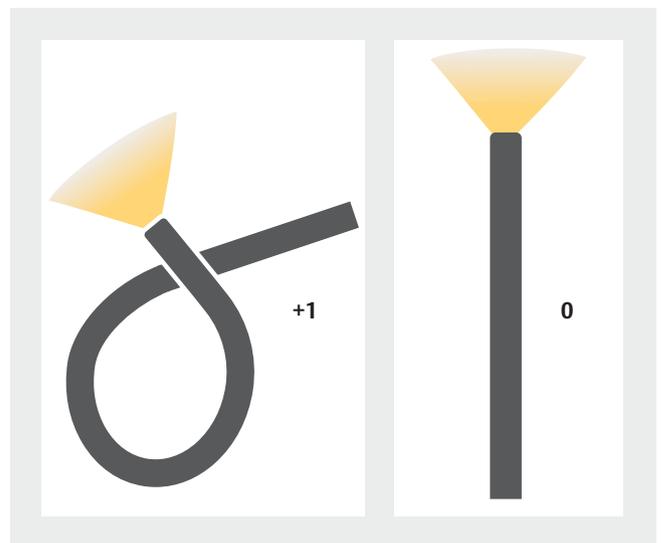
greater distance between the coils from one-time point to another (► **Fig. 3**).

Shaft movement without tip progression: Ideally the whole colonoscope moves concurrently, but this is not always the case. This measure refers to the situation where the tip of the colonoscope stands still but the rest is moving and potentially stretching the colon and thus mesenteric tissue. A movement cannot be considered concurrent if the tip of the colonoscope is lodged while the shaft is moving (a low score). A concurrent movement is when the tip of the colonoscope and the shafts is moving equally (a high score). The analysis is a comparison of the movement differences between the tip and the shaft (the scale has no physical unit) (► **Fig. 4**).

Looping: The occurrence of loops during a colonoscopy is correlated to patient discomfort, pain or even incompleting procedure [18]. A quantified measure of the amount of looping during a procedure has never been investigated. In 2002 Rogen et al. introduced an automatic classification method of protein structure, called Writhe [19]. Writhe is a statistic for how knotted a protein is and we applied the same method to describe the colonoscope. The quantification of scope crossings provides us with a score for how much the colonoscope is crossing its own path in a 3D space during the procedure (► **Fig. 5**). The scale has no physical unit, is time-independent, and goes from 0 to 1. A score of 0 represents a completely straight colonoscope whereas a score of 1 represents a complete loop in either direction. The total score is the area under the curve where the Y axis is the Writhe value, and the X axis is time. A threshold for meaningful looping was set at above 0.5 as we want to measure meaningful loops and not general bending of the colonoscope.



► **Fig. 4** Shaft movement without tip progression is a measurement of movement differences between the tip and the shaft. Increasing score reflects a greater difference between the movements of the tip compared to the shaft. The score has no physical unit.



► **Fig. 5** Looping is a measurement of how looped the colonoscope is during a colonoscopy. A score of 0 reflects a completely straight colonoscope (right side) and a score of 1 reflects a bent colonoscope overlapping itself (left side). The score has no physical unit.

Furthermore, only the first 70 cm of the colonoscope was analyzed to avoid including bends and loops outside the phantom model in the score.

Gathering validity evidence for the newly developed assessment tool: We gathered evidence of validity according to Messick's contemporary framework and its five sources: content, response process, internal structure, relationship to other variables, and consequences [20]. These categories of evidence need to be collected to support the construct validity of inferences.

► **Table 1** Demographics.

Group	12 novices	12 experienced	P value
Mean age (range)	29.5 (25–32)	55.5 (41–66)	<0.001
Sex (Female/Male)	F = 8 (67%) / M = 4 (33%)	F = 1 (8%) / M = 11 (92%)	0.002
Mean years since graduation (range)	1.9 (0–8)	28.5 (16–39)	<0.001
Mean colonoscopies performed (range)	7.4 (0–34)	6,863 (2,000–10,000)	<0.001
Mean colonoscopies per year (range)	7.4 (0–34)	512 (300–1000)	<0.001
Mean gastroscopies performed (range)	48 (0–200)	NA	NA
Simulator experience (yes/no)	10/2	12/0	0.166

The measurements are meant to measure the core technical skills of colonoscopy. Collecting content evidence for the measurements was based on a previously validated assessment tool and phantom model [17, 21], and the opinions of an expert panel consisting of colonoscopists and engineers. Response process was ensured by standardized introduction to the equipment and the phantom model, and all data-collection were recorded in a uniform file-format. When using multiple items in a test that is intended to measure the same thing (technical performance), a goal is to know to what extent the items measure the same thing (reliability of the test). Internal consistency was explored using Cronbach's alpha. We assumed that the data reflected a probability distribution and hence enable us to test the assumption that the 3D-CoPS is correlated to experience, i. e. relationship to other variable (using Pearson's *r*). To ensure technical competence based on 3D-CoPS Contrasting Groups methods were used (Consequence) to calculate a pass/fail.

Participants and equipment

Twenty-four volunteer physicians were included and divided into two groups based on clinical experience, defined by the colonoscopy volume; 12 novices with less than 50 colonoscopies and 12 experienced with more than 140 colonoscopies. A standardized introduction lasting 30 minutes was given to novices without former experience. All participants were given written and oral information and signed a consent form before entering the study. Demographics are shown in ► **Table 1**.

The study was done using a setup with an Olympus colonoscope (CF-H180DL, Evis Exera II video center CV-180, Olympus Medical System Ltd, Tokyo, Japan), the ScopeGuide (UPD-3, Olympus, Tokyo, Japan), and a phantom model (Kyoto Kagaku Colonoscopy Training Model, by, Japan). The ScopeGuide was connected to an Olympus receiver box (UCES-3, Olympus, Tokyo, Japan) and an Intel NUC computer device (NUC7i3BNK, Intel Corporation, Santa Clara, California, United States) running Windows 10.

The training model contains a 130-cm long rubber colon which can be configured into different scenarios. Each participant performed twice on the same case (case 3) with an alpha loop formation in the sigmoid colon. A maximum of 10 minutes was allowed to reach the cecum. Data collection (XYZ-coordinates) was started at the insertion of the colonoscope into the

anus and ended when the participant had a clear view of the cap representing the cecum. No feedback was given during or in between the performances. The data was independent of the UPD-3 unit settings and saved on the Intel NUC computer device. The five measurements and 3D-CoPS were developed in Python 3.7 and the XYZ dataset were applied for calculations of the measurements for each participant.

Sample size

The sample size was calculated based on data from a previous trial [17]. A significance level of 5% and a power of 0.9 required a minimum of 6 participants in each group.

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 clinical-trials.gov with trial identification number NCT03401723.

Statistics

Statistical analysis was done in IBM SPSS statistics (PASW, version 22; SPSS Inc, Chicago, Illinois, United States) and STATA software version 14.0 (College Station, TX). The level of statistical significance was set at $\alpha < .05$ for all tests. Internal consistency was evaluated using Cronbach's alpha. To allow comparison between the five measurements and the two groups the scores were standardized by converting to z-score for each performance. Consequentially, a z-score of 1 was equivalent to one standard deviation higher than the mean score. The coefficient of variation was calculated as the standard deviation divided by the mean. The correlation between performances in the first and second attempt (i. e. test-retest reliability) was explored using Pearson's *r* [22]. Comparison of the mean scores of the two performances between groups was made by using an independent sample t-test for continuous data and the Pearson chi-square test for categorical data.

► **Table 2** Mean score for the five measurements and cecal intubation time.

	Novices Mean (SD)	Experienced Mean (SD)	P
Travel length (cm)	737 (397)	378 (155)	<.001
Tip progression (cm)	3418 (2073)	1525 (944)	<.001
Chase efficiency (meter per second)	250 (69)	188 (46)	.001
Shaft movement without tip progression (no physical unit)	0.65 (0.12)	0.79 (0.11)	<.001
Looping (no physical unit)	617 (622)	223 (165)	.006
Cecal intubation time (minutes)	4:35 (2:39)	1:35 (0:45)	<.001

Results

Data collection was conducted from October 2017 to February 2018. Twelve experienced and 12 novices participated. One participant in the experienced group was excluded due to a large and unintended movement of the phantom which affected the whole data-set. All novices but one reached the cecum twice within the time limit of 10 minutes.

Each of the five measurements was analyzed individually; statistically significant differences in performance were demonstrated between the novices and experienced participants in all measurements (► **Table 2**).

Internal consistency of the 3D-CoPS was high: Cronbach's alpha=0.78. Indicating that all five outcome measures were highly correlated, i.e. measured the same construct.

3D-colonoscopy progression score: Mean scores of the five measurements, based on the two performances, were calculated based on the standardized z-scores. The measurements were weighted equally (20%) and combined to a single score of progression, the 3D-CoPS. The novices scored significantly lower than the experienced; -0.454 (SD 0.707) and 0.495 (SD 0.303), $P<0.001$, respectively (► **Fig. 6**).

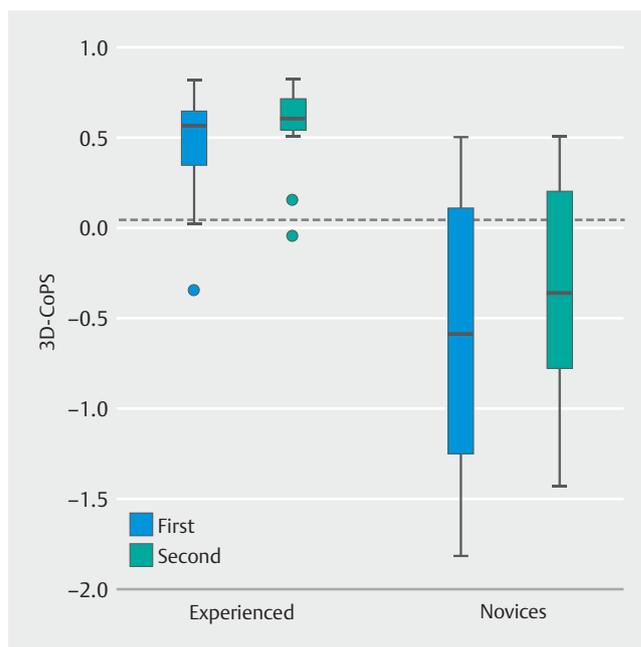
A Pearson correlation was done revealing a strong test-retest reliability between the two tests ($r=0.86$, $P<0.001$) (► **Fig. 7**).

A pass/fail was set based on a contrasting groups method (► **Fig. 8**).

Based on the passing score one experienced failed both tests and one experienced failed the first try. Three novices passed both tests and the rest failed both the first and second test. The 3D-CoPS and clinical experience had a moderate correlation ($r=0.61$, $P<0.001$).

Discussion

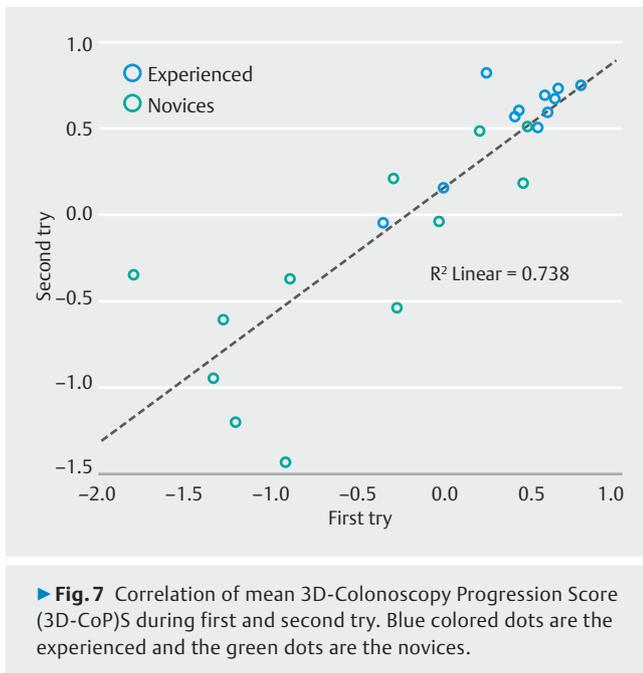
We developed a computerized assessment tool, the 3D-CoPS, to assess technical competence in simulated colonoscopy. Five different measures were created based on XYZ-coordinates from along the length of the colonoscope: 1. Travel length; 2. Tip progression; 3. Chase efficiency; 4. shaft movement without tip progression; and 5. Looping. The 3D-CoPS and the results on each of the five measurements revealed a statistically significant difference between groups in performance. Internal



► **Fig. 6** Box-Plot of 3D-Colonoscopy Progression Score (3D-CoPS) showing outliers (*) and passing score resembled by the dotted line. Blue colored box is the first try and the green box is the second try.

consistency and reliability of the results were good, and evidence of validity was established.

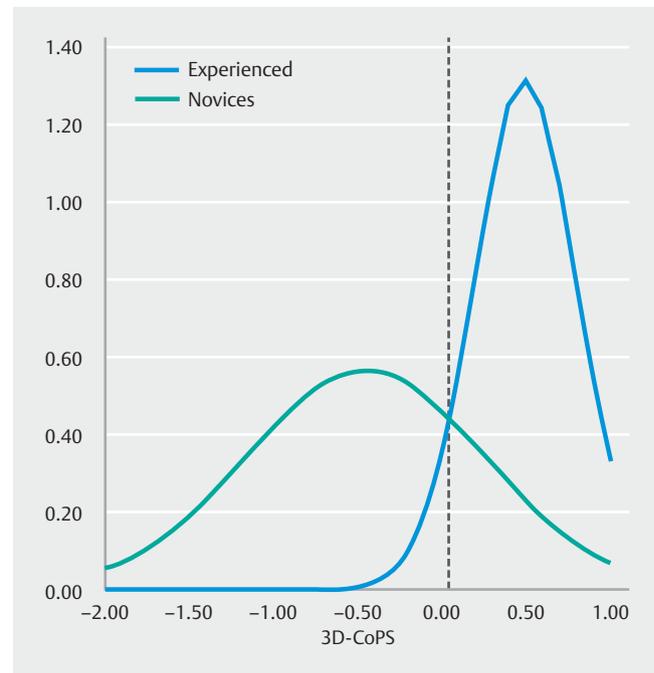
Technical skills are often evaluated by using single measures (e.g. direct observation of procedural skills or cecal intubation rate). Safe and smooth steering of the colonoscope relies on more than just one technical aspect during insertion, hence combining several technical measures increase reliability and reduce possible bias [23]. By combining the five different measurements into a single score, the 3D-CoPS, we reduce the weight of each measurement, and therefore reduce misinterpretation based on single parameters. Unexpectedly, even for the experienced endoscopists the travel length from anus to cecum was 2.9 times the length of the rubber colon and for novice endoscopists, 5.6 times. This is not in line with previous reports on the distance from the anus to the cecum (computed tomography [CT] scan: 189 cm (range, 75–257) and colonoscopy:



83cm (range, 49–150)) [24, 25]. Results from CT scan are closer to the real length of the rectum and colon, compared with the length measured during colonoscopy. During a colonoscopy the sigmoid colon tends to be shortened during de-looping maneuvers, hence the length may be underreported. The traveled distance is not the same as the shortest distance from A to B. This indicates that even for an experienced endoscopist there is room for improvement to optimize the distanced traveled.

Tip progression is the cumulative distance from the tip to the optimal path line from anus to the cecum. A high score indicates an uneven route during insertion. Although some degree of stretching of the colon is inevitable as the instrument pushes inward, with subtle movements, the endoscopist may be able to achieve almost “direct” passage with minimal stretch and bends, hence minimizing the distance from the tip to the optimal path line. Similar motor skills have been described in the DOPS and Mayo Colonoscopy Skills Assessment Tool (MCSAT) as tip control and safe endoscope advancement techniques. However, the assessment is based on the amount of verbal guidance or hands-on assistance during a procedure, making it prone to observer bias. Furthermore, the assessment tools are designed to evaluate a training program’s ability to meet specific requirements rather than assessing a single procedure [10, 26, 27].

Pain during a colonoscopy is most likely to occur in the sigmoid colon, and 80% of reports are due to loops or straightening the colon [28]. To traverse the colon as gently and as rapidly as possible, the operator needs to keep the shaft as straight as possible, avoid losing “one-to-one” movements and avoid over-angulation of the tip [29]. Experienced endoscopists often try to avoid situations where the tip becomes lodged and further attempts to advance the colonoscope leads to the shaft stretching the colon wall. Shaft movement without tip progres-



sion and looping are supposed to resemble situations where the colonoscope stretches the colon wall and potentially induces discomfort, pain, or even risk. Novices endoscopists, who may not know the techniques needed to prevent unwanted events may instead forcefully push the colonoscope creating loops or shaft movements without tip progression.

The looping scores indicate that experienced endoscopists had a less bent colonoscope and spent less time with loop formation during insertion. This is in line with previous studies investigating loop management in a simulation-based setup [21, 30]. In a clinical perspective the time spent from anus to cecum on the phantom was very short even for those with experience (► **Table 2**) [31, 32]. Moreover, in the clinic patients resemble a much more heterogeneous group with various anatomic formations and loop management in the clinic pose a greater challenge compared to our setup [32, 33].

Direct observation assessment tools assess numerous domains of the colonoscopy procedure one of which is technical aspects, such as pace, tip control and scope handling [12, 27]. These are essential aspects when evaluating the trainee during a procedure but are based on a subjective interpretation of the assessment tools and are therefore subject to rater bias [16, 34]. The simulated environment offers an opportunity to deconstruct the colonoscopy procedure into its parts based on the five different measurements. This allows trainees to identify and comprehend the important procedural steps and train single aspects to enhance performance.

The major endoscopy societies around the world recommend regular assessment of endoscopists in general. Direct observation by a supervisor assessing the trainee and procedure-

related quality measures, such as cecal intubation rate, time spent and complications is frequently used [35]. Experts within the field of endoscopy have been calling not only for a means for objective assessment but also for a continuous evaluation to follow a learning curve as trainees progress. The assessment tools are prone to bias and are resource demanding, which reduces compliance regarding their continued use. Sedlack et al. followed all gastroenterology fellows in training at the Mayo institution during a 3-year study. The overall reported compliance rate with completing the assessment tool on each procedure was 62% but was initially as low as 21% [27]. Preisler et al. explored the correlation between CoPS and patient experienced discomfort but had to exclude 29% of recordings [36]. Excluding user errors, the compliance was as high as 89% (67/75) favoring automatic assessment.

A meta-analysis investigating the effect and type of feedback in simulation-based training found that feedback increases performance [37]. Traditionally, supervisors are needed to provide feedback when training, but supervisors remain a scarce resource, which is the reason why self-practice is gaining ground in the simulation centers. Continuous and automatic assessment as typically seen in the virtual reality simulators provides the trainees with easily accessible feedback and access to the learning process. In the clinic, however, continuous assessments remain difficult without supervisors and the workload needed decreases the compliance, hence the number of procedures assessed for meaningful learning curves becomes inaccessible [38]. The 3D-CoPS circumvent these problems being automatic, unbiased and feasible with the possibility of assessing, giving feedback and access to learning curves, in both a simulation and a clinical environment.

Limitations

First of all, assessing colonoscopy is much more than just technical performance and 3D-CoPS do not assess technical skills such as mucosa visualization and the ability to handle therapeutic tools. Furthermore, non-technical skills, such as pathology identification and the management of patient discomfort have been acknowledged as essential components [39]. Until recently Olympus was the only manufacturer with a magnetic endoscopic imaging system; accordingly, 3D-CoPS has been tested only on Olympus equipment. The study was conducted in a simulation-based setup that offered a standardized and feasible model for gathering initial validity evidence. A possible limitation was the missing resemblance to the clinic, but a meta-analysis is investigating the use of simulation-based assessments found a positive correlation with patient-related outcomes [40]. However, clinical studies should be performed to explore the validity and usefulness of the system in colonoscopies on patients.

We used a relatively small sample of physicians; nevertheless, we found significant differences in each of the five measurements and the 3D-CoPS. Validation studies have been criticized for the expert-novice comparison [41]. When comparing a proficient group (e.g., experts) with one that is not (e.g., true beginners such as medical students), a large difference in competence is to be expected. The novice group in this study was

made up of physicians with various experience, ensuring a very high baseline capability compared with other studies within the same field, leading to a stricter pass/fail standard than if all novices were “true beginners.” Inclusion criterion for the novices was physicians having done fewer than 50 clinical colonoscopies; only four novices had no clinical experience. One of whom had already participated in simulation-based training programs and had passed the tests in colonoscopy and gastroscopy. The rest had mixed experienced with up to 200 clinical gastroscopies and 34 clinical colonoscopies. This might explain why some novices performed so well and why the pass/fail standard was so high.

The selected case was straightforward with an alpha loop in the sigmoid colon. The ease of the case may have interfered with the discriminative ability between the more experienced physicians by introducing a ceiling effect to the assessment tool. Conversely, increasing the level of difficulty would have increased the risk of the novices not being able to complete the case. During a colonoscopy, the endoscopist might need to change the position of the patient to advance the colonoscope. In our study, the phantom model was fixed in the supine position. A position change of the phantom model shifts all the coils approximately 90 to 180 degrees in either direction and thus affects travel length, tip progression, and chase efficiency. Shaft movement without tip progression is less likely to be affected because the measurement is the difference in movement between the tip and the shaft, whereas looping remains completely unaffected of external movement. A position change may unintentionally favor or disfavor our measurement scores since the position change could be meaningful for further advancement during a procedure. As a result, position change during a procedure requires mathematical adaption in all our measurements to account for the change in the dataset. Ongoing clinical studies are dealing with these issues.

In theory, 3D-CoPS is fully implementable in the clinic and can continuously assess endoscopists. Furthermore, 3D-CoPS could add to existing training programs providing the trainee and the supervisors' with information on technical skills and learning curves [14, 42]. Moreover, the tool could be used logistically to optimize time schedules for patients in control programs. A patient with a previous colonoscopy now has a score of procedural difficulty, this information can be used in the planning to ensure the technical competence of the endoscopist, and also the time needed to complete the colonoscopy. Quality measures for gastrointestinal endoscopy units have been stated to constitute high-quality endoscopy. Many of these procedure-related indicators demand retrospective registrations and system developments for data tracking and interpretation. Furthermore, efforts needed to gather the relevant information have led to a low response rate [4]. Computerized automatic assessment of every colonoscopy could provide the units with a procedure specific quality indicator based on their entire production.

Conclusion

In conclusion, the study presents a novel, real-time computerized assessment tool for colonoscopy with strong evidence of validity based on Messicks framework. XYZ-coordinates from coils along the length of the colonoscope were sampled, five different technical measurements were developed and built into a combined score of progression, the 3D-CoPS. With further development, 3D-CoPS could provide feedback for trainees, aid in the certification process, and help ensure competent performance of colonoscopies.

Acknowledgements

The authors would like to thank Dr. Paul Frost Clementsen, PhD, DMSc, for his illustrations explaining the scores.

Competing interests

The authors declare that they have no conflict of interest.

References

- [1] Benson VS, Patnick J, Davies AK et al. Colorectal cancer screening: A comparison of 35 initiatives in 17 countries. *Int J Cancer* 2008; 122: 1357–1367
- [2] Rembacken B, Hassan C, Riemann JF et al. Quality in screening colonoscopy: position statement of the European Society of Gastrointestinal Endoscopy (ESGE). *Endoscopy* 2012; 44: 957–968
- [3] 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* 2017; 49: 1175–1189
- [4] Day LW, Cohen J, Greenwald D et al. Quality indicators for gastrointestinal endoscopy units. *VideoGIE* 2017; 2: 119–140
- [5] Valori R, Senore C, Thomas de Lange I et al. Performance measures for endoscopy services: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative ESGE Quality Improvement Committee (QIC) Endoscopy Service Working Group. *Endoscopy* 2018; 50: 1186–1204
- [6] Rex DK, Boland CR, Dominitz JA et al. Colorectal Cancer Screening: Recommendations for Physicians and Patients from the U.S. Multi-Society Task Force on Colorectal Cancer. *Am J Gastroenterol* 2017; 112: 1016–1030
- [7] Chak A, Cooper GS, Blades EW et al. Prospective assessment of colonoscopic intubation skills in trainees. *Gastrointest Endosc* 1996; 44: 54–57
- [8] Sedlack RE. Training to competency in colonoscopy: Assessing and defining competency standards. *Gastrointest Endosc* 2011; 74: 355–366.e2
- [9] Baruk JH, Cohen ER, Feinglass J et al. Residents' procedural experience does not ensure competence: a research synthesis. *J Grad Med Educ* 2017; 9: 201–208
- [10] Sedlack RE, Coyle WJ, Obstein KL et al. ASGE's assessment of competency in endoscopy evaluation tools for colonoscopy and EGD. *Gastrointest Endosc* 2014; 79: 1–7
- [11] Walsh CM, Ling SC, Khanna N et al. Gastrointestinal Endoscopy Competency Assessment Tool: reliability and validity evidence. *Gastrointest Endosc* 2015; 81: 1417–1424.e2
- [12] Vassiliou MC, Kaneva PA, Poulou BK et al. Global assessment of gastrointestinal endoscopic skills (GAGES): A valid measurement tool for technical skills in flexible endoscopy. *Surg Endosc* 2010; 24: 1834–1841
- [13] Siau K, Dunckley P, Anderson J et al. Changes in scoring of direct observation of procedural skills (DOPS) forms in endoscopy training and their impact on competence assessment. *Trainees' Sect Symp #FOAMed Film Fest Educ Dragons Den 2017: A11.1–A11*
- [14] Sedlack RE, Coyle WJ. Assessment of competency in endoscopy: Establishing and validating generalizable competency benchmarks for colonoscopy. *Gastrointest Endosc* 2016; 83: 516–523.e1
- [15] Coe SG, Panjala C, Heckman MG et al. Quality in colonoscopy reporting: An assessment of compliance and performance improvement. *Dig Liver Dis* 2012; 44: 660–664
- [16] Konge L, Vilmann P, Clementsen P et al. Reliable and valid assessment of competence in endoscopic ultrasonography and fine-needle aspiration for mediastinal staging of non-small cell lung cancer. *Endoscopy* 2012; 44: 928–933
- [17] Nerup N, Preisler L, Svendsen MBS et al. Assessment of colonoscopy by use of magnetic endoscopic imaging: Design and validation of an automated tool. *Gastrointest Endosc* 2015; 81: 548–554
- [18] Mark-Christensen A, Brandsborg S, Iversen LH. Magnetic endoscopic imaging as an adjuvant to elective colonoscopy: a systematic review and meta-analysis of randomized controlled trials. *Endoscopy* 2015; 47: 251–261
- [19] Rogen P, Fain B. Automatic classification of protein structure by using Gauss integrals. *Proc Natl Acad Sci U S A* 2003; 100: 119–124
- [20] Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: Theory and application. *Am J Med* 2006; 119: 166.e7–16
- [21] Plooy AM, Hill A, Horswill MS et al. Construct validation of a physical model colonoscopy simulator. *Gastrointest Endosc* 2012; 76: 144–150
- [22] Downing S, Yudkowsky R. *Assessment in Health Professions Education*. New York: Routledge; 2009
- [23] Shah SG, Thomas-Gibson S, Brooker JC et al. Use of video and magnetic endoscope imaging for rating competence at colonoscopy: Validation of a measurement tool. *Gastrointest Endosc* 2002; 56: 568–573
- [24] Teshima CW, Zepeda-Gómez S, AlShankiti SH et al. Magnetic imaging-assisted colonoscopy vs conventional colonoscopy: A randomized controlled trial. *World J Gastroenterol* 2014; 20: 13178–13184
- [25] Duncan JE, McNally MP, Sweeney WB et al. CT colonography predictably overestimates colonic length and distance to polyps compared with optical colonoscopy. *Am J Roentgenol* 2009; 193: 1291–1295
- [26] 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
- [27] Sedlack RE. The Mayo Colonoscopy Skills Assessment Tool: Validation of a unique instrument to assess colonoscopy skills in trainees. *Gastrointest Endosc* 2010; 72: 1125–1133.e3
- [28] Shah SG, Brooker JC, Thapar C et al. Patient pain during colonoscopy: An analysis using real-time magnetic endoscope imaging. *Endoscopy* 2002; 34: 435–440
- [29] Haycock A, Cohen J, Saunders BP et al. *Cotton and Williams' Practical Gastrointestinal Endoscopy: The Fundamentals*. Oxford: Wiley Blackwell; 2014
- [30] 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

- [31] Cadoni S, Gallittu P, Sanna S et al. A two-center randomized controlled trial of water-aided colonoscopy versus air insufflation colonoscopy. *Endoscopy* 2014; 46: 212–218
- [32] Asai S, Fujimoto N, Tanoue K et al. Water immersion colonoscopy facilitates straight passage of the colonoscope through the sigmoid colon without loop formation: Randomized controlled trial. *Dig Endosc* 2015; 27: 345–353
- [33] Haycock A, Koch AD, Familiari P et al. Training and transfer of colonoscopy skills: a multinational, randomized, blinded, controlled trial of simulator versus bedside training. *Gastrointest Endosc* 2010; 71: 298–307
- [34] Sedlack RE, Shami VM, Adler DG et al. Colonoscopy core curriculum. *Gastrointest Endosc* 2012; 76: 482–490
- [35] 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
- [36] Preisler L, Bulut M, Svendsen MS et al. An automatic measure of progression during colonoscopy correlates to patient experienced pain. *Scand J Gastroenterol* 2018; 53: 345–349
- [37] Hatala R, Cook DA, Zendejas B et al. Feedback for simulation-based procedural skills training: A meta-analysis and critical narrative synthesis. *Adv Heal Sci Educ* 2014; 19: 251–272
- [38] Chung JII, Kim N, Um MS et al. Learning curves for colonoscopy: A prospective evaluation of gastroenterology fellows at a single center. *Gut Liver* 2010; 4: 31–35
- [39] Hitchins CR, Metzner M, Edworthy J et al. Non-technical skills and gastrointestinal endoscopy: a review of the literature. *Frontline Gastroenterol* 2018; 9: 129–134
- [40] Brydges R, Hatala R, Zendejas B et al. Linking simulation-based educational assessments and patient-related outcomes. *Acad Med* 2015; 90: 246–256
- [41] Stefanidis D, Sevdalis N, Paige J et al. Simulation in surgery: what's needed next? *Ann Surg* 2015; 261: 846–853
- [42] Zupanc CM, Burgess-Limerick R, Hill A et al. A competency framework for colonoscopy training derived from cognitive task analysis techniques and expert review. *BMC Med Educ* 2015; 15: 1–11