Teaching Colonoscopy

from Novice to Competence

Arjun Koch

Teaching Colonoscopy

from Novice to Competence

Arjun Koch

Financial support for printing this thesis was kindly given by Abbott B.V., Covidien, Dr. Falk Pharma Benelux B.V., ERBE Nederland B.V., Ferring B.V., Olympus Nederland B.V., Tramedico B.V., Zambon Nederland B.V. and the department of Gastroenterology and Hepatology of the Erasmus MC – University Medical Center of Rotterdam.

ISBN : 978-94-6169-519-2

Lay-out and printing: Optima Grafische Communicatie, Rotterdam.

Teaching Colonoscopy – from Novice to Competence

Colonoscopie - van kennis gestuurd naar competentie gericht onderwijs

PROEFSCHRIFT
ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus
Prof.dr. H.A.P. Pols
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
vrijdag 27 juni 2014 om 09:30 uur

door

Arjun Dave Koch

geboren te Harlingen

ERASMUS UNIVERSITEIT ROTTERDAM

PROMOTIECOMMISSIE

Promotor:	Prof.dr. E.J. Kuipers
Overige leden:	Prof.dr. M.J. Bruno
	Prof.dr. P.D. Siersema
	Prof.dr. J.J.B. van Lanschot
Copromotoren:	Dr. R.A. de Man
	Dr. E.J. Schoon

TABLE OF CONTENTS

Chapter 1

Training and competence assessment in gastrointestinal endoscopy: a systematic review

Vivian E. Ekkelenkamp*, Arjun D. Koch*, Robert A. de Man, Ernst J. Kuipers

**Authors contributed equally to this work*

Department of Gastroenterology and Hepatology, Erasmus MC - University Medical Center, Rotterdam, the Netherlands

Submitted

ABSTRACT

introduction

Training procedural skills in gastrointestinal endoscopy once focused on threshold numbers. However, as threshold numbers poorly reflect individual competence, the focus gradually shifts towards a more individual approach. Tools to assess and document individual learning progress are being developed and incorporated in dedicated training curricula. However, there is a lack of consensus and training guidelines differ worldwide, which reflects uncertainties on optimal set-up of a training program.

aims

The primary aim of this systematic review was to evaluate the currently available literature for the use of training and assessment methods in GI endoscopy. Secondly, we aimed to identify the role of simulator-based training as well as the value of continuous competence assessment in patient-based training. Thirdly, we aimed to propose a structured training curriculum based on the presented evidence.

methods

A literature search was carried out in the available medical and educational literature databases. The results were systematically reviewed and studies were included using a predefined protocol with independent assessment by two reviewers and a final consensus round.

results

The literature search yielded 5846 studies. Ninety-four relevant studies on simulators, assessment methods, learning curves and training programs for gastrointestinal endoscopy met the inclusion criteria. Twenty-seven studies on simulator validation were included. Good validity was demonstrated for four simulators. Twenty-three studies reported on simulator training and learning curves, including 17 RCT's. Increased performance on a virtual reality simulator was shown in all studies. Improved performance in patient-based assessment was demonstrated in 14 studies. Four studies reported on the use of simulators for assessment of competence levels. Simulator-based performance did not reflect competence in patient-based endoscopy. Eight out of fourteen studies on colonoscopy, ERCP and EUS reported on learning curves in patient-based endoscopy and proved the value of this approach for measuring performance. Ten studies explored the numbers needed to gain competence, but the proposed thresholds varied widely between them. Five out of nine studies describing the development and evaluation of assessment tools for gastrointestinal endoscopy provided insight in performance of endoscopists. Five out of seven studies proved that intense training programs result in good performance.

conclusions

The use of validated virtual reality simulators in the early training setting accelerates learning of practical skills. Learning curves are valuable for continuous assessment of performance and are more relevant than threshold numbers. Future research will strengthen these conclusions by evaluating simulation-based as well as patient-based training in gastrointestinal endoscopy. A complete curriculum with assessment of competence throughout training needs to be developed for all gastrointestinal endoscopy procedures.

INTRODUCTION

The focus on training in procedural skills in gastrointestinal (GI) endoscopy is shifting from threshold numbers towards an individual approach. This illustrates the awareness that the classic master-apprentice model may not reflect all necessary aspects of training. Moreover, the old adage 'see one, do one' seems no longer appropriate for educating health professionals to perform complex technical procedures, such as flexible endoscopy.[1] Virtual reality (VR) simulators may be of benefit in the education of gastroenterology trainees. However, a substantial part of training still has to be patient-based. The assessment of a trainee's competence is not clearly defined and competence benchmarks for trainees are sparse. The use of threshold numbers is nowadays considered a poor surrogate marker for competence. Keeping track of one's performance by measuring skill development seems preferable. However, training guidelines differ worldwide and there is no consensus on the skills a trainee has to possess at the end of education. On top of that, for most procedural skills in flexible endoscopy, the proper assessment tools to measure these skills are lacking.

The aim of this systematic review was therefore to evaluate the available literature on different training and assessment methods in gastrointestinal endoscopy. Secondly, we aimed to identify the role of simulator training and competence development in patient-based training, specifically for procedures that normally will be learnt during residency. Thirdly, we aimed to propose a structured training curriculum based on the presented evidence.

METHODS

literature search strategy

A systematic literature search was carried out in July 2013 in seven different medical and educational literature databases: Embase, Medline OvidSP, Web of Science, Cochrane central, Google Scholar, Research and Development Resource Base (RDRB) and Education Recourse Information Center (ERIC). There was no restriction regarding time of publication or language. The search strategy for Medline OvidSP is shown in supplementary file 1.

in- and exclusion criteria

All studies pertaining to training and assessment in gastrointestinal endoscopy (colonoscopy, endoscopic retrograde cholangiopancreatography (ERCP), endosonography (EUS), and upper gastrointestinal endoscopy) were included in this review. The studies were to report outcome measures with respect to learning curves, assessment methods or tools and training programs including simulators. Two reviewers independently examined all retrieved studies. When disagreement existed over studies to be in- or excluded, these were discussed until consensus was reached. Reviews, systematic reviews, meta-analyses and abstracts were excluded, as well as studies on tools to improve completion of colonoscopy. However, reference lists of potential relevant systematic reviews and meta-analyses were checked for any missed papers.

data extraction and analysis

For each study, the methods, way of assessment, and endpoints were recorded according to a predefined protocol. Two reviewers extracted all data. The quality of the studies was appraised and the reviewers assigned a level of evidence to each study using a tool developed by the Oxford Centre for Evidence-based Medicine (CEBM) as shown in Table 1.[2] A grade of recommendation was given for each subgroup of studies included in this systematic review using the same tool provided by the CEBM (Table 1). The validation method and type of each simulator study was designated according to the consensus guidelines for validation of virtual reality simulators as described by Carter et al. [3] Validation of simulators is in most cases performed by demonstrating different types of validity. Validity in itself is defined as the extent to which an assessment tool, in this case a simulator, measures what it is supposed to measure. One of the simplest forms of validity is face validity. This is demonstrated by questioning a defined group of subjects, to judge the simulator on realism between the simulator and the real activity. Usually, a group of experts is questioned. This is why the term expert validity is also used. Construct validity describes the extent to which the simulator can distinguish between different levels of expertise. The most used method of establishing construct validity is that the simulator can distinguish beginners from more experienced endoscopists and experts by the simulators performance parameters. Reliability of the simulator relates to the power of the simulator to provide consistent results. The most commonly used test is the test-retest reproducibility. It predicts to what extent a subject can 'beat the test' by repeated assessment. The most powerful evidence of validity is concurrent validity. This refers to the level of which performance on the simulator correlate to the real activity, in this case patient-based endoscopy.

Since we aimed to provide a complete overview of the available literature on training and assessment in GI endoscopy, the included studies were fairly heterogeneous. Therefore, it was judged that statistical pooling of the data was not suitable.

Table 1. *Levels of evidence and grades of recommendation (CEBM) of the reported studies.*

Level of evidence	
1a	Systematic reviews (meta-analysis) containing at least some trials of Level Ib evidence, in which results of separate, independently conducted trials are consistent
1b	RCT of good quality and of adequate sample size (power calculation), cohort study with good follow up
2a	RCT of reasonable quality and/or if inadequate sample size
2b	Non-randomized trials, comparative research parallel cohort
3	Non-randomized, noncomparative trials, descriptive research
4/5	Case series, expert opinions, including the opinion of work group members
Grades of recommendation	
A	Consistent level 1 studies
B	Consistent level 2 or 3 studies *or* extrapolations form level 1 studies
C	Level 4 studies *or* extrapolations from level 2 or 3 studies
D	Level 5 evidence *or* troubingly inconsistent or inconclusive studies of any level

RESULTS

inclusion

Figure 1 shows the flowchart of the selection process of the included studies. Ninety-four studies investigating simulators, assessment methods, learning curves and training programs for gastrointestinal endoscopy were included in this review. In order to provide a systematic overview of these studies, they were divided into different categories (simulator

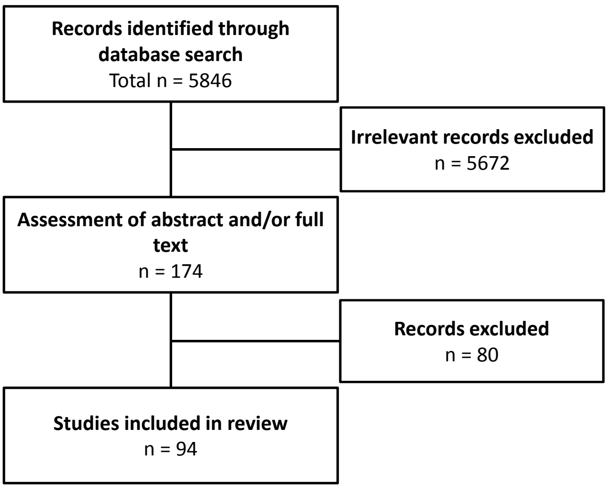

Figure 1. Flow diagram of the studies included

training, learning curves, numbers needed to gain competence, assessment of performance and evaluation of (patient-based) training models). In the more detailed discussion, we will focus on studies providing level 1 and 2 evidence.

simulator validation studies

Twenty-seven simulator validation studies were retrieved. [4-29] We included eleven studies on colonoscopy, two on flexible sigmoidoscopy, three studies on basic flexible endoscopy in general, one study on EGD, six studies on ERCP, two on EUS, and one study on dexterity exercises in forward viewing endoscopy. All studies are shown in Table 2. Besides the ERCP and EUS studies, all five other categories of studies focused on conventional, forward viewing flexible endoscopy with a large overlap in outcome parameters. Procedures like ERCP and EUS show profound differences compared to basic forward viewing flexible endoscopy, not only because of combination with radiological or ultrasonographical imaging but also because of a complete different perception by the endoscopist in side viewing endoscopy. We have therefore analysed them separately from the larger group that we refer to as forward viewing flexible endoscopy procedures. Eight validation studies on flexible endoscopy tasks were performed using the Simbionix GI Mentor VR computer simulator. [8-10, 12, 14, 15, 22, 25] Two studies reported on face validity. The largest study included 35 experts and demonstrated good face validity for colonoscopy. [15] A smaller study reported low level of realism as judged by six experts on all modules of the simulator. [14] All studies reported consistent results and good construct for performance metrics on procedure times of the GI Mentor. These procedural times varied from time to cercal intubation, time spent with clear view and time spent with endoscope loops. Although these types of parameters, measuring a time aspect, are usually considered surrogate markers for competence, it seems to be the most consistent and therefore the most reliable parameter to distinguish between competence levels. There is a fairly large heterogeneity on other outcome parameters. Five studies reported on the AccuTouch Immersion Medical computer simulator. [4, 19-21, 27] Two studies reported on face validity with conflicting results. Again, as for the GI Mentor, realism was judged as valid by experts for the colonoscopy module but not for the complete set of modules on the simulator as a whole. The AccuTouch simulator seemed to have the same construct validity profile as the GI Mentor. That is, construct validity was consistently reported as good for performance measures related to procedural times in all published studies. Three validation studies reported on the Olympus Endo TS-1 VR computer simulator for colonoscopy. [13, 16, 28] Face validity was rated as good by two studies and all three demonstrated good construct validity on all studied procedures. One study reported good construct validity of the Kyoto Kagaku Colonoscope Training Model. Face validity was not studied. [24] The last study demonstrated good face, construct and concurrent validity in a bovine explant colon model. [26]

Six validation studies were performed on ERCP. Two studies were feasibility studies and no formal validation was done. These two were both in mechanical models. [11, 23] Only one val-

Table 2. *Studies on validation of simulators for all gastrointestinal endoscopy procedures.*

Study	Setting and participants	Simulator	Procedure	Methods	Validation method	Results	LOE
Bhutani[6] (1998)	Feasiblity study for EUS	Live porcine model	EUS	Comparing EUS images to autopsy findings	Feasibility only, no validation	Visualization of most anatomical landmarks in a swine model seem comparable to human anatomy	4
Bhutani[5] (2000)	Feasiblity study for interventional EUS	Live porcine model	EUS	Creation, location and puncture of pseudo-lesions	Feasibility only, no validation	Demonstrated feasibility to perform FNA of a pseudo-lymph node, submucosal lesion, pancreatic tissue and celiac neurolysis	4
Neumann[23] (2000)	Feasibility study for ERCP	Erlangen Endo-Trainer	ERCP	ERCP cannulation, sphincterotomy and stent placement	Feasibility only, no validation	Demonstrated feasibility to perform ERCP with deep cannulation, sphincterotomy and stent placement	4
Datta[4] (2002)	15 Novices, 15 intermediate and 15 experts	AccuTouch	Sigmoidoscopy	3 cases performed by all after a familiarization tour. Recorded parameters were mucosa visualized, time taken, scope pathlength, time in red-out and efficiency ratio.	Construct validity	Significant difference of mucosa visualized and efficiency ratio between all groups, difference in time and scope pathlength for novices compared to experienced	2b
Ferlitsch[10] (2002)	13 Novices, 11 experts	GI Mentor	Flexible endoscopy	VR simulator tasks on the GI Mentor included EndoBubble, gastroscopy and colon-oscopy. Additionally, 7 novices performed unsupervised training: 2 hours, 5 days a week for 3 weeks	Construct validity and early learning curve	Baseline assessment revealed significant differences in all tasks in favor of experts proving construct validity. After 3 weeks of training, differences were no longer seen whereas these differences persisted for the no-training group	2b
MacDonald[19] (2003)	10 Novices, 19 experienced and 5 experts	AccuTouch	Sigmoidoscopy	3 VR sigmoidoscopies. Performance is measured by the simulator	Construct validity	Significant differences on all parameters between the untrained and trained endoscopists. Experts perform better than senior residents but with no significant differences.	2b
Mahmood[20] (2003)	11 Novices (<10 procedures), 7 intermediates (11-100 procs), 7 experienced endoscopists (>100 procs)	AccuTouch	Colonoscopy	2 VR colonoscopies on the simulator. Performance is measured by the simulator	Construct validity	Significantly improved performance with increased experience on all parameters measured except path-length	2b

Table 2. *Continued*

Study	Setting and participants	Simulator	Procedure	Methods	Validation method	Results	LOE
Ritter[25] (2003)	9 Novices and 5 intermediate experienced endoscopists (100-250 procs)	GI Mentor	Dexterity	3 VR tasks EndoBubble level 2 on the GI Mentor II simulator. Performance parameters were recorded for each attempt and compared to previous attempts.	Construct validity and reliability	Intermediates performed better in each trial without improvement of performance. Novices improved significantly in completion time. Consistency and reliability were greater than 0.8 for both groups on all measures except completion time in the novice group	2b
Sedlack[27] (2003)	6 Novices, 6 trainees (88-202 colonoscopies) and 10 experts	AccuTouch	Colonoscopy	2 VR colonoscopies on the simulator. Performance is measured by the simulator. A survey on realism by experts	Face and construct validity	Favourable degree of realism as judged by experts but substantially easier. Only procedure time, insertion time and time in red-out significantly different between groups	2b
Sedlack[30] (2003)	10 Trainees and 10 trainers	Live porcine model, Erlangen Endo-Trainer, GI Mentor	ERCP	procedures in live porcine models, the Erlangen Endo-Trainer and the GI Mentor. Survey on realism and performance. Ranking of results	Face validity	Overall results for realism and educational value highest for the Erlangen model and lowest for the GI Mentor. GI Mentor was easiest incorporated in a training program	2b
Moorthy[22] (2004)	11 Novices, 11 intermediates (10-50 procs), 10 experienced (>200 procs)	GI Mentor	EGD	2x2 VR tasks by all participants. Performance measured by simulator and video assessment by blinded experts using standardized assessment form	Construct validity and reliability	Significant differences on all simulator output parameters between the 3 groups. Significant differences in global scores between groups, only percentage of mucosa visualized correlated with global scores	2b
Felsher[9] (2005)	39 Novices and 36 experienced endoscopists	GI Mentor	Colonoscopy	Colonoscopy task on the GI Mentor II including insertion, withdrawal and polypectomy	Construct validity	Experienced ednoscopists performed the tasks more efficient, visualized more mucosa, achieved greater polypectomy rates and spent more time in clear view of the lumen than novices ($P<0.05$)	2b
Grantcharov[12] (2005)	10 Novices, 10 intermediate and 8 experienced endoscopists	GI Mentor	Colonoscopy	Single VR colonoscopy on the GI Mentor II simulator	Construct validity	Significant differences in performance by experts versus intermediates and novices on all recorded parameters. Differences between intermediates and novices not significant	2b

Table 2. *Continued*

Study	Setting and participants	Simulator	Procedure	Methods	Validation method	Results	LOE
Leung[17] (2007)	6 Unexperienced (<100 ERCPs), 9 experienced (>100 ERCPs), 8 experts	ERCP Mechanical Simulator - EMS	ERCP	Performance times are measured for ERCP total procedure times and simulated fluoroscopy times. Participants completed a survey on realism	Face and construct validity	Significant shorter procedure times for more experienced endoscopists on all aspects. Very high credibility scores	2b
Sedlack[26] (2007)	13 Novices, 13 trainees (100-150 colonoscopies) and 13 experts	Ex vivo bovine colon model	Colonoscopy	Single colonoscopy on a bovine ex vivo colon model. Standardized performance measures were maximum insertion depth and time, scope length, withdrawal time, % of mucosa visualized and quality of mucosal exam. Performance was compared to known patient-based performance for each endoscopist	Face, construct and concurrent validity	Favourable degree of realism as judged by experts, significantly greater performance compared to less experienced groups. High correlation to patient-based cecal intubation times (r=0.764)	2b
Frimberger[11] (2008)	26 Trainees and 7 trainers	X-Vision ERCP Training System	ERCP	ERCP cannulation, sphincterotomy and stent placement	Face validity	Favourable results in all categories assessed	4
Koch[16] (2008)	26 Novices, 23 experts	Endo TS-1	Colonoscopy	VR tasks included a hand-eye coordination game and a colonoscopy. Experts completed a survey on realism	Face and construct validity	Experts completed tasks significantly faster and rated face validity as good.	2b
Koch[15] (2008)	35 Novices, 15 intermediates, 20 experienced and 35 experts	GI Mentor	Colonoscopy	VR tasks included a hand-eye coordination game and 2 colonoscopies. Experts completed a survey on realism	Face and construct validity	Novices performed significantly worse than the other groups. The experienced groups performed equal. Expert validity was rated as good	2b
Haycock[13] (2009)	10 Novices, 13 intermediates and 11 experts	Endo TS-1	Colonoscopy	3 VR colonoscopy cases on the Endo TS-1 with increasing difficulty	Face and construct validity	Experts performed significantly better compared to intermediates and novices. Face validity was rated as good	2b
von Delius[29] (2009)	6 Advanced, 10 intermediate, 12 novice endoscopists	X-Vision ERCP Training System	ERCP		Construct validity, Expert validity, didactic value	Significant differences for construct validity between experts, experienced and novice endoscopists	2b

Table 2. *Continued*

Study	Setting and participants	Simulator	Procedure	Methods	Validation method	Results	LOE
Bittner[7] (2010)	6 Novices and 6 experts	GI Mentor	ERCP	2 ERCP tasks performed on the simulator	Face and construct validity	Significant differences between novices and experts performing simulated ERCP cases demonstrating both face and construct validity	2b
Fayez[8] (2010)	12 Novices (<5 procedures), 8 experienced (>50 procedures)	GI Mentor	Colonoscopy	3 VR colonoscopy cases on the GI Mentor II with increasing difficulty	Construct validity	Differences in performance between novices and experienced on the simulator are greater in the more complex scenarios. During easy tasks only time to cecum differed. In difficult tasks ther difference in experience resulted in significantly better performance for the experienced group	2b
Kim[14] (2010)	5 Novices, 6 experts	GI Mentor	Flexible endoscopy	Novices perform 50 supervised cases of which 6 predetermined cases are used for evaluation, experts perform 6 cases. Evaluation form on realism by experts	Face and construct validity	No significant overall differences between 2 groups on performance parameters. Only differences on time parameters of the procedures. No realistic representation	2b
Leung[18] (2011)	22 Experienced endoscopists	EMS and ex vivo porcine model	ERCP	Same ERCP tasks; scope insertion, selective bile duct cannulation and biliary stent placement, were performed on the EMS and the PSM. Survey on credibility was performed before and after traning	Face validity	Before hands-on practice, both EMS and PSM received high scores. After practice, there was a significantly greater increase in confidence score for EMS than PSM ($p<0.003$). Participants found EMS more useful for training ($p=0.017$)	2b
McConnell[21] (2012)	5 Novices and 6 experts	AccuTouch	Flexible endoscopy	18 VR colonoscopies and 6 gastroscopies performed by all participants. Survey on realism by all	Face and construct validity	No significant overall differences between 2 groups on 38/57 performance parameters. Only differences in 19/57 parameters of which 8/19 on time parameters of the procedures. No realistic representation as rated by experts	2b
Plooy[24] (2012)	18 Novices and 21 experienced endoscopists	Kyoto Kagaku Colonoscope Training Model	Colonoscopy	2x4 standard cases on the simulator by all participants. Performance measures included completion rate, time to cecum and peak force	Construct validity	Experts had significantly higher completion rates and shorter times to cecum for all cases. Peak force was significantly lower in 2 cases	2b

Table 2. *Continued*

Study	Setting and participants	Simulator	Procedure	Methods	Validation method	Results	LOE
Sugden[28] (2012)	30 Novices, 10 intermediates (100-500 colonoscopies) and 10 experienced (>500 procs)	Endo TS-1	Colonoscopy	VR tasks included 3 abstract tasks, 9 part-procedural and 1 whole procedure task. Each set was performed 2 times by intermediates and experienced and 10 times by novices. Benchmarks were established and learning curves for novices to reach these	Construct validity	Construct validity was demonstrated for 3 abstract, 4 part-procedural and the whole procedure task. Significant differences were demonstrated between novices and both experienced groups but not between intermediates and experienced. Learning curves plateaued consistently at the 9th or more attempt	2b

idation study was performed using a VR computer based simulator. This study demonstrated both face and construct validity for the ERCP modules in the Simbionix GI Mentor II simulator. [7] A similar study was done for the X-Vision ERCP Training System, a mechanical simulator, showing both face and construct validity. [29] Two studies on the same mechanical ERCP training simulator were performed by the same research group. [17, 18] The ERCP Mechanical Simulator (EMS) demonstrated a good construct validity and excellent face validity. In a direct comparison to an ex vivo porcine stomach model, the EMS was rated more realistic and useful. Another study compared live porcine models versus the Erlangen Endo-Trainer versus the Simbionix GI Mentor VR simulator for ERCP. [30] The Erlangen model scored highest on realism and educational value. The GI Mentor scored lowest. However, it was felt that the GI Mentor was more easily incorporated in a training program. Although the validation studies for ERCP simulation comprised a fairly heterogeneous group of simulators, the strongest evidence was provided for the mechanical simulators. For EUS, only two studies by the same author reported on feasibility to perform EUS and FNA in a porcine model. [5, 6] No attempt at validation has been published to date.

simulator training and learning curve studies

Twenty-three studies reported on simulator training and learning curves. [31-53] Twenty studies reported on forward viewing flexible endoscopy (3 EGD, 3 sigmoidoscopy, 14 colonoscopy), one study reported on EUS, one on ERCP and one on training haemostasis in upper GI bleeds. The studies are shown in Table 3. Eleven studies were performed using the AccuTouch Immersion Medical VR computer simulator for training, ten with a level 2 evidence and one study with a level 1 evidence. [31, 33, 34, 40, 44, 45, 47, 49-52] All studies on flexible sigmoidoscopy and colonoscopy had a randomized design and compared simulator-based training groups versus controls. Acquired competence was evaluated using the same simulator and in six studies also during patient-based assessment. The most consistent outcome parameters demonstrating improved performance were on procedural times, caecal intubation rates, and times in red-out, meaning that luminal view was lost. Patient comfort scores were measured in two studies. [40, 50] One study favoured simulator training versus no simulator training prior to starting patient endoscopies, the second study showed no difference between groups.

Six studies were carried out using the Simbionix GI Mentor VR simulator for training and learning curves. [32, 34, 36-38, 41] Five studies provided level 2 evidence, one study level 1 evidence. Four studies were on colonoscopy tasks, two on EGD. All studies demonstrated that simulator training improved performance of novices. There were no learning effects for experienced endoscopists. Due to the heterogeneity of these studies, improved performance could not be expressed in terms of exact numbers. Performance was assessed by means of the simulator construct in three studies. Two studies used patient-based assessment for evaluation of the simulator-based learning effect. The competence parameters that con-

Table 3. *Simulator-based training and learning curve studies.*

Study	Setting and participants	Simulator	Procedure	Methods	Assessment method	Results	LOE
Tuggy[31] (1998)	Single center, randomized, 5 novice endoscopists in the simulator group, 5 controls	Accutouch	Sigmoidoscopy	5 And 10 hours VR sigmoidoscopy training on the simulator versus no training	Improvement of performance during patient based sigmoidoscopy	Significantly shorter procedure times, no difference in times in red-out or completeness of mucosal inspection	2b
Gerson[40] (2003)	Single center, randomized, 9 novices in simulator-trained group, 7 in control group	Accutouch	Sigmoidoscopy	Unsupervised training on the simulator versus patient-based training	5 Patient-based sigmoidoscopies per trainee evaluated by supervisor and patients	Patient-based trained group completed significantly more cases, completed moe cases independently, performed more successful retroflexion. There was no difference in procedure times and patients satisfaction or discomfort	2a
Eversbusch[38] (2004)	Single center, randomized, 2 groups; group I: 8 experienced endoscopists (>200 procedures), 10 intermediate (<50 procedures), 10 novices. Group II: 10 novices psychomotor training versus 10 novices control.	GI Mentor	Dexterity and colonoscopy	Group I: learning curve psychomotor training using EndoBubble task 10 times. Group II: VR colonoscopy followed by psychomotor training versus no training and again VR colonoscopy	Improved performance during VR colonoscopy	No learning curve during psychomotor training for experts, short learning curve for intermediate experienced and a longer learning curve for novices. Statistically significant improved performance during VR colonoscopy on all parameters after psychomotor training versus control.	2b
Giulio[41] (2004)	Multicenter, randomized, 22 novices	GI Mentor	EGD	10 hours unsupervised training on the simulator versus no training	20 Patient-based gastroscopies per trainee evaluated by non-blinded supervisors	Significantly more complete examinations, less need for assistance and more positive scores by supervisors in the simulator trained group	2a
Mahmood[47] (2004)	Single center, 26 Novices performing VR colonoscopies on the AccuTouch Immersion Medical simulator	Accutouch	Colonoscopy	38 trials of 5 consecutive VR colon-oscopies. Performance measures are recorded by the simulator	Improvement of performance measures on the simulator in the abcense of feedback	In the abcense of feedback there was no learning effect measured on the simulator after repetitive training	2b

Table 3. *Continued*

Study	Setting and participants	Simulator	Procedure	Methods	Assessment method	Results	LOE
Sedlack[49] (2004)	Single center, randomized, 4 novices to VR colonoscopy training versus 4 controls	Accutouch	Colonoscopy	VR colonoscopy training for 6 hours	Patient-based assessment of colonoscopy performance using standardized assessment forms by experts and patient survey on discomfort	Study group outperformed controls on all parameters except insertion time during first 15 procedures and on 3 parameters (depth of insertion, independent completion and ability to identify landmarks) after 30 colonoscopies (P<0.05)	2a
Sedlack[50] (2004)	Single center, randomized, 19 novices versus 19 controls	Accutouch	Sigmoidoscopy	19 Novices receiving 3 hours of simulator-based training prior to patient-based sigmoidoscopy versus 19 controls	1 Week of patient-based sigmoidoscopy. Assessment of performance using a standardized assessment form by supervisors and patients survey on discomfort	Median patient discomfort scores were significantly lower in the study group (P<0.01). No differences in procedural performance as judged by supervisors	2a
Ahlberg[34] (2005)	Single center, randomized, 12 novice colonoscopists with only gastroscopy experience	GI Mentor	Colonoscopy	VR colonoscopy training to a predefined expert level on the simulator versus no training	Patient-based colonoscopy	Significantly improved performance in the simulator-trained group (P=0.0011) and 4.53 times more likely to reach the cecum	2a
Cohen[37] (2006)	Multicenter, randomized, 45 novices	GI Mentor	Colonoscopy	Randomized to no simulator training versus 10 hours of unsupervised simulator training	Assessment of technical and cognitive performance as well as patient discomfort during the first 200 consecutive patient-based colonoscopies	Simulator group had significantly higher objective competence rates during the first 80 cases. The overall number needed to reach a 90% competence level was 160 cases in both groups	1b
Barthet[35] (2007)	17 novice trainees performing EUS	Live porcine model	EUS	Visualizing anatomical landmarks, perform FNA and celiac neurolysis	Pre- and post-test assessment of performance	Increased performance after training in both diagnostic (visualization) and therapeutic procedures (time to complete and precision, not the rate of technical errors)	3
Buzink[36] (2007)	Single center, 35 novices and 5 expert endoscopists	GI Mentor	Dexterity and colonoscopy	All participants perfoming a predefined protocol colonoscopy and hand-eye coordination tasks	Assessment by simulator construct repeating predefined tasks on the simulator itself	Novices significantly improve performance in the early learning curve by training on the GI Mentor II, whereas experts do not	2b

Table 3. *Continued*

Study	Setting and participants	Simulator	Procedure	Methods	Assessment method	Results	LOE
Maiss[48] (2007)	27 Novices randomized to intensive simulator training versus controls	CompactEASIE Simulator	Haemostasis	14 Novices are trained 12 times in a 7 months period in endoscopic skills, injection therapy, clipping and bandligation	Final assessment by blinded experts	Study group performed significantly better on all tasks as judged by blinded experts. Results were comparable to similar studies performed in New York and France.	2a
Park[52] (2007)	Single center, randomized, 12 novices receiving simulator-based training versus 12 controls	Accutouch	Colonoscopy	All participants performed baseline VR colonoscopy. SG received 2-3 hours of training on the simulator	Patient-based assessment by blinded experts using GRS. Scores were compared to simulator performance metrics	SG had a significantly higher GRS. Only 2/8 simulator performance metrics correlated with GRS (procedure time and time in red-out)	2a
Ferlitsch[39] (2010)	Single center, randomized, 28 novices, 14 received simulator training prior to patient-based gastroscopy training	GI Mentor	EGD	5 to 20 hours of VR simulator training in upper GI endoscopy and training games with the first 2 hours supervised	Patient-based gastroscopy with standardized assessment of performance by the unblinded supervising expert on time taken, technical and diagnostic accuracy and blinded patients evaluation on discomfort and pain.	Time and technical accuracy was significantly better in the simulator-trained group ($P<0.05$) in the early learning curve, after 60 procedures time was still significantly different. Diagnostic accuracy and patient discomfort or pain was not different in any point of time.	2a
Haycock[42] (2010)	Multicenter, randomized, blinded, 36 novice colonoscopists	Endo TS-1	Colonoscopy	16 hours VR colonoscopy training versus patient-based training	Patient-based and simulator-based assessment	No differences in performance between subjects and controls during patient-based assessment. Subjects performed significantly better in simulator-based assessment.	1b
Kruglikova[44] (2010)	Single center, randomized, 22 novices	Accutouch	Colonoscopy	Simulator training with concurrent feedback versus no feedback	Post-training simulator-based assessment	Subjects reached proficiency levels significantly faster than controls. This effect diminished during the retention/transfer tests	2a
Kruglikova[45] (2010)	Single center, 30 female novices, 10 nurses, 10 endoscopy nurses, 10 residents	Accutouch	Colonoscopy	10 repetitions of same VR task on the simulator	Comparing first and last colonoscopy task to each group and expert reference	Residents performed the tasks faster with a similar learning curve pattern, other parameters were equal.	2b

Table 3. *Continued*

Study	Setting and participants	Simulator	Procedure	Methods	Assessment method	Results	LOE
Snyder[51] (2010)	Single center, randomized, 8 proctored versus 5 unproctored trainees	Accutouch	Colonoscopy	Both groups performed VR colonoscopy assessment: baseline directly after training & after 4.5 months	A combined proficiency score of 10 (best performance on all parameters) was the maximum score and considered proficient	All participants improved to proficient level post-training. Retention scores after 4.5 months were unchanged for all participants	2a
Kaltenbach[43] (2011)	Single center, pilot study, 3 trainees	Mechanical colon model	Colonoscopy	Half-day one-on-one training using a colon model with ScopeGuide imaging creating different loops during the trainees colonoscopy learning curves	Colonoscopy performance was appraised (non-blinded) before and after the training	Performance scores and CIRs improved significantly	4
Lim[46] (2011)	Multicenter, randomized, 16 trainees in ERCP	ERCP Mechanical Simulator - EMS	ERCP	2 Sessions of hands-on training in selective bile duct cannulation on the EMS for the study group	Patient-based ERCP with cannulation success rates, cannulation times and blinded supervisor assessment	Significant higher cannulation success rates in less time in the study group. Same competence scores for both groups	1b
Van Sickle[32] (2011)	Multicenter, prospective, 41 novice endoscopists	GI Mentor	Dexterity	Novice endoscopists training EndoBubble on the simulator	Pre- and posttraining assessment using Colonoscopy case module I outcome parameters and GAGES assessment form	Performance on the simulator, as measured by the simulator parameters, improved significantly after performing EndoBubble tasks ranging 13 +/-10 for EndoBubble level 1 and 23 +/- 16 for level 2	2b
Ende[53] (2012)	Multicenter, randomized, 28 novices, EGD	Various	EGD	3 groups; clinical+simulator training, clinical training only, simulator training only. Different simulators used.	Assessment by simulator and blinded experts using standardized forms	Clinical trained groups outperformed the simulator trained group significantly. Results in the combination group were better than clinical training alone	1b
Ahad[33] (2013)	Single center, randomized, 32 novice endoscopists	Accutouch	Colonoscopy	Low versus high-fidelity model simulator training	Pre- and posttraining assessment using colonoscopy tasks on the Immersion Accutouch simulator	Colonoscopy skills acquisition of basic endoscopy skills on a low-fidelity model is just as effective as on a high-fidelity simulator	1b

sistently improved significantly were; (I) procedure time, (II) caecal intubation rate (CIR): a direct comparison of simulator-training versus controls showed a 4.5-fold increased caecal intubation rate in the simulator-training group in the early learning curve [34], (III) time with clear view, (IV) time of endoscope looping, and (V) objective performance scores, as judged by expert supervisors during patient-based endoscopy assessment. Improved performance in the simulator-trained groups versus controls was observed in up to 60 patient-based assessed EGDs and 80 procedures in colonoscopy training. Only one study used the Olympus Endo TS-1 colonoscopy simulator for training. [42] This multicenter, randomized study compared simulator-based training versus patient-based training. Blinded experts assessed performance during patient-based endoscopy. Both groups showed equal performance. (LOE 1b) One multicenter, randomized study was performed using all kinds of simulators. [53] The study showed that patient-based training with complementary simulator training was superior to patient or simulator-based training alone. (LOE 1b) One study was done on ERCP. [46] This study had a multicenter, randomized design. It demonstrated significantly higher cannulation success rates in less time in the study group after training on the ERCP Mechanical Simulator. (LOE 1b) One study was performed evaluating the CompactEASIE simulator, a mechanical simulator with an ex vivo porcine stomach. [48] Significant improvement in skills in endoscopic haemostatic therapy was demonstrated with a sufficient level of evidence. No previous formal validation of the model was carried out. Only one study was performed on the subject of learning diagnostic and therapeutic EUS. [35] Only a description of improved performance on live porcine models before and after a hands-on training course was provided. No formal statistical calculation was carried out. The model had not been previously validated.

simulator competence assessment studies

Four studies reported on the use of simulators for assessment of competence. [54-57] Two studies focussed on colonoscopy, one on sigmoidoscopy and one on both EGD and sigmoidoscopy. The studies are summarized in Table 4. Only two studies reached a 2b LOE. [55, 57] In both studies performance parameters derived from the simulators did not correlate to performance scores given by blinded experts. It seems that current simulators lack the discriminative power to assess performance and determine competence levels in patient-based endoscopy.

Table 4. *Simulator competence assessment studies.*

Study	Setting and participants	Simulator	Procedure	Assessment method	Results	LOE
Moorthy[55] (2004)	Single center, 7 novices, 7 intermediates (20-80 procs), 6 experts (>200 procs) performing VR sigmoidoscopy on the AccuTouch Immesion Medical simulator	AccuTouch	Sigmoidoscopy	Dual video assessment by blinded experts using standardized assessment forms and simulator performance parameters	Significant discrimination of skills levels by the global scores. Only "time in red-out" correlated with global scores and not for time, depth of insertion and mucosa visualized	2b
Phitayakorn[56] (2009)	23 Experts perform a single colonoscopy task on the GI mentor II	GI Mentor	Colonoscopy	Performance parameters results with their range are demonstrated as an attempt to set a benchmark	Considerable variety in performance on the simulator among experts	4
Sarker[57] (2010)	37 Trainees (<100 procs) and 18 consultants (>500 procs) performing both patient-based and simulator-based EGD and sigmoidoscopy, GI Mentor	GI Mentor	EGD and sigmoidoscopy	Assessment of patient-based endoscopy by blinded experts using video recordings, simulator-based assessment only by composed outcome measures of the simulator itself: mucosal score and time score	Construct validity was good for the patient-based assessment ($P=0.000$-0.002) but not for the simulator based outcome parameters ($P=0.263$-0.701)	2b
Elvevi[54] (2012)	Single center, 12 novices in colonoscopy, GI Mentor	GI Mentor	Colonoscopy	Competence level assessment at the start of training and after 60 patient-based colonoscopies using 2 GI Mentor colonoscopy modules	Only time to cecal intubation improved significantly during simulator assessment. The simulator is considered unuseful for competence assessment	3b

learning curves

Fifteen studies reported on learning curves for colonoscopy (n=8), ERCP (n=5) and EUS (n=2). [58-72] These are shown in Table 5 (A,B,C, respectively).

For colonoscopy, four studies reached a sufficient evidence level 1 or 2. [59, 61, 62, 65] These studies had a prospective design and evaluated 8 to 41 trainees with procedure numbers varying from 2887 to 4351. However, outcome measures and use of competence standards were fairly heterogeneous. The studies reported on caecal intubation rate (CIR) or completion rate, time to cecum, or a combination of those outcomes. One group described the learning curve by means of scoring different aspects of the procedure on a newly developed assessment tool (Mayo Colonoscopy Skills Assessment Tool), but also described learning curves for outcomes such as CIR. [65] The number of colonoscopies that trainees needed to perform in order to achieve a CIR of > 85-90% varied from 150 to 280 procedures.

From the five studies focusing on ERCP, only two reached a sufficient evidence level. [60, 71] These described a prospective evaluation of respectively 17 and 20 trainees, with the following outcome measures: subjective score regarding performance (overall and per part of the procedure) on a 6-point scale where a score of 1,2 or 3 was considered competent, and success of selective cannulation of the common bile duct (CBD) or pancreatic duct (PD). One study concluded that an overall sufficient score was reached after 137 (probability of success = 0.8) or 185 ERCPs (probability of success = 0.9). [60] A different group reported that an 85% selective cannulation rate was reached after 70 procedures for the PD and after >100 ERCPs for the CBD. [71]

The two studies on EUS described the performance per anatomic station of the procedure. [70, 72] There was a large variability in achieving overall competence, with acceptable performance after a range of 255 to >400 EUS procedures.[70] One study did not report on overall competence, but stated that 78 procedures were necessary for competence in duodenal examination. [72]

Table 5A. *Learning curves and threshold numbers as a measure for EGD, sigmoidoscopy and colonoscopy.*

Study	Setting and participants	n participants	Assessment method – outcome measures	Results	LOE
Sedlack [65] (2011)	Single center, prospective, trainees at different levels of experience	41	Learning curves through assessment (MCSAT) with scores from 1-4: competence when score >3.5. CIR & time to cecum.	4103 of 6635 colonoscopies (62%) assessed (mean n per trainee: 399; 95% CI 365-433). Competence after mean 275 procedures. CIR of 85% after 250-275 colonoscopies, time to cecum <16 min after 275.	1b
Koch[61] (2012)	*Single center, prospective, trainees at different levels of experience*	19	Learning curves through self-assessment (Rotterdam Assessment Form for Colonoscopy). CIR and time to cecum.	2887 colonoscopies (152 per trainee, range 91-347). Baseline CIR: 65% improved to 78% after 100 and 85% after 200 colonoscopies ($p<0.001$). After 280 ≥90% CIR. Time to cecum decreased from baseline 13:10 min to 8:30 min after 200.	1b
Chung[59] (2009)	Single center, prospective, GI fellows with no colonoscopy experience	12	Learning curves per 50 procedures for adjusted completion rate (>90%) and time to cecum (<20 min).	3243 colonoscopies (n per trainee not mentioned). First 50 completion was 37% and improved to 94% after 250. Mean time improved from 12.9 min to 8.0 min. After 200 procedures, >90% completion rate within 20 min was reached.	2b
Lee[62] (2008)	Multicenter, prospective, first-year GI fellows with no colon-oscopy experience	24	Outcome: adjusted CIR (>90%) and time to cecum (<20 min). Learning curves per 50 colonoscopies.	4351 colonoscopies (n per trainee not mentioned). CIR per 50 consecutive procedures: 71.5% - 82.6% - 91.3% - 94.4% - 98.4% - 98.7% ($p<0.05$). Time to cecum decreased from 11.2 min to 6.6 min. Competence after 150 procedures.	2b
Tassios[67] (1999)	Single center, prospective, trainees with no colonoscopy experience	8	Learning curves for completion rate. Outcome: completion of colonoscopy (CIR) per training year and per number of procedures performed.	N = 978 colonoscopies performed by trainees (median n 1st year: 43, range 31-69 and 2nd year 91 (79-143). Completion rates in %: first-year 33 (range 15-42); second-year 60 (range 55-79) and third-year 75 (range 58-94) ($p<0.001$). Per number: after 60 52.6 (95% CI 45.2-59.9); after 100 67.1 (95% CI 59.1-75.2) and after 180 76.8 (95% CI 66.1-87.6).	3
Marshall[63] (1995)	Multicenter, prospective, first and second-year trainees with different levels of experience.*	9	Frequency of reaching cecum <30 min in last 7 months of 1st and 2nd year training.	N colonoscopies: 186 (n per trainee not mentioned) by first-year and 203 by second-year fellows. CIR 1st year: 53.7%, 2nd year 85.8%. Time to cecum 28,19 and 9 min, respectively. Difference in performance by groups ($p<0.001$).	3

Table 5A. *Continued*

Study	Setting and participants	n participants	Assessment method – outcome measures	Results	LOE
Selvasekar[66] (2012)	Single center, prospective, fellows with different levels of experience in colonoscopy	6	Learning curves by completion and time to cecum. Cusum analysis. Competence defined as significant reduction in time and 80% CIR within 35 min.	1498 out of 2904 colonoscopies performed by fellows (mean n per fellow 249). Mean procedure time 30.2 +-15 min; decreased significantly during first 120 procs (p<0.001). 80% completion within 35 min after 114 colonoscopies.	3
Parry[64] (1990)	Single center, retrospective, trainee with no experience in colonoscopy	1	Learning curve by cusum analysis for completion rate.	305 of 334 consecutive colonoscopies analyzed. First 100 procedures completion rate 67%, 101-200 was 88% and 201-305 91%. After 200 procedures, steady >90% completion.	4
Cass[73] (1993)**	Single center, prospective crosssectional design, surgical and GI trainees with different levels of endoscopy experience. *EGD*	12	Intubation esophagus	Median 113 UGI endoscopies per trainee (range 54-162). Esophagus intubation 90% after 50 procedures, but declined and reached 80% after 100 procedures.	2b
Cass[73] (1993)**	*Colonoscopy*	12	Intubation cecum	Median 49 colonoscopies per trainee (range 39-127). CIR of 80% after 50 procedures but only 85% at 100 procs.	2b
Chak[74] (1996)	Single center, prospective, GI fellows at various stages of endoscopy training, attendings	12 fellows, 17 attendings	Cecal intubation rate, time to cecum per training year, assistance needed. Adequacy of threshold of 100 colonoscopies	496 colonoscopies performed (79 by 5 first-years, 102 by 7 second-years, 315 by attendings). Second-years performed mean of 123 colonoscopies prior to study. First-years required assistance in 92% of procs vs 36% for second-years. CIR second-years vs attendings: 84 vs 94%, p<0.05). Time to cecum second-years vs attendings: 14.5 vs 10.5 min (p<0.01). 90% success rate within 15 min not reached after 100.	2b
Vassiliou[80] (2010)**	Multicenter, prospective design, surgical and GI trainees with different levels of endoscopy experience. Single observations. *UGI endoscopy*	86	Score on GAGES for UGI, 3 different groups: <35 cases (1), 35-130 cases (2), >130 cases (3)	86 evaluations, group (1) n=35, (2) n=22, (3) n=29. Mean +-SD score: (1) 14.4 +- 3.7; (2) 17.8 +- 1.8; (3) 19.1 +- 1.1. Difference between groups significant (p<0.05) but not for (2) and (3). Plateau score at n=50.	2b

Table 5A. *Continued*

Study	Setting and participants	n participants	Assessment method – outcome measures	Results	LOE
Vassiliou[80] (2010)**	*Colonoscopy*	57	Scores on GAGES colonoscopy, 2 groups: cutoff <50 cases (1) and <140 cases (2); novices and experienced	57 evaluations, cutoff (1) 29 novices vs 28 experienced. Mean +- SD score: (1) novices 11.8 +-3.8 vs. experienced 18.8 +- 1.3 (p<0.001). Cutoff (2) 32 vs. 25. Novices 12.4 +- 4.2 vs. experienced 18.8 +- 1.3. No difference between cutoff groups. Plateau score at n=75.	2b
Spier[79] (2010)	Single center, surgical trainees who participated in 2-month endoscopy rotation, survey	21	Numbers performed, CIR, perception of colonoscopy training	100% response rate, 15 2nd year residents, 6 4th year. Mean 80 +- 35 colonoscopies during rotation. Average CIR 47% (range 9-78%). Adequacy of training: 67% of 2nd year felt it was adequate vs 100% 4th years.	3
Spier[78] (2010)	Single center, retrospective, GI fellows at various stages of endoscopy training	11	Total colonoscopy time, time to cecal intubation, independent completion rates. Adequacy of threshold of 140 colonoscopies	770 colonoscopies performed (369 by 9 first-years, 158 by 4 second-years, 243 by 5 third-years). Improvements from first- to third-year: mean colonoscopy time from 48 to 33 min, time to cecal intubation 19 to 11 min, completion rate 63 to 92% (all p<0.001). No independent completion of >90% after 140 colonoscopies.	3
Leyden[77] (2011)	Single center, retrospective, GI and surgical trainees with >2 years endoscopic experience	13	Completion rates, polyp detection rates, withdrawal time in subset of patients, comparison of GI and surgical trainees	1998 and 1081 colonoscopies performed by GI and surgical trainees, respectively. Crude completion rate for GI vs surgical: 84 vs 78% (p<0.0001). PDR was 21 vs. 14% (p<0.001); withdrawal time 5 vs 2.5 min (p=0.003).	3
Church[75] (2002)	Single center, prospective, trainees with different endoscopy experience	18	Performance during first 125 colon-oscopies, completion rate defined as cecum reached by trainee as a percentage of completion by staff. Time to cecum	Completion improved from first 25 cases to fifth 25 cases: 43.1% to 75.1%. Time to cecum from 18.7 to 17.1 min.	4
Hawes[76] (1986)	*Single center, prospective, residents rotating on GI ward performing sigmoidoscopy*	25	Assessment of overall skill on 6-point competence scale (1-3 not competent, 4-6 competent), accuracy of diagnosis	495 of 662 sigmoidoscopies were graded. Initial 10 examinations largely graded 3 or less. Examinations 10 to 25: largely grade 3 or 4, and >25, 82% was graded 4 or above. With increasing competence score, more correct diagnoses.	3

*2 testing periods. Period 1: 4 first-year fellows, 2 second-year. Period 2: 3 first-year fellows, 4 second-year.

** These studies reported on both UGI endoscopy and colonoscopy, the study population is the same.

Table 5B. *Learning curves and threshold numbers as a measure of performance in ERCP.*

Study	Setting and participants	n participants	Assessment method – outcome measures	Results	LOE
Jowell[60] (1996)	Single center, prospective, fellows with different levels of ERCP experience	17	Outcome: grading of performance on 6-point scale. Competence: score 1,2,3; overall and per part of procedure (i.e. CBD cannulation). Learning curve by probability of success.	1450 of 1796 ERCPs were evaluated. Median n per trainee=132 (range 57-186). Overall score: 0.8 after 137 ERCPs and 0.9 after 185 procedures. Score and numbers for CBD cannulation: 0.65 after 180 ERCPs.	2b
Watkins[71] (1996)	Multicenter, prospective, trainees with no previous ERCP experience	20	Outcome: selective cannulation of CBD and PD. Analysis per 10 consecutive procedures.	641 ERCPs performed by trainees (mean 31, range 10-96). First 10 ERCPs (n=199); PD cannulation 46%, CBD cannulation 32%. Rapid increase during first 40 procedures; 85% selective cannulation after 70 ERCPs for PD and >100 for CBD.	2b
Waller[69] (2009)	*Single center, prospective, trainee without ERCP experience*	1	Outcome: successful cannulation. Learning curves through cumulative failure charting and (cusum) analysis (unacceptable failure rate 0.35, acceptable failure rate 0.20).	290 ERCPs. Cumulative failure charting: after 60 ERCPs no unacceptable performance. Acceptable performance after 100 procedures with cusum analysis.	3
Biau[58] (2008)	Single center, retrospective, endoscopist without ERCP experience	1	Learning curves through cusum analysis, success rates for selective cannulation (acceptable failure rate 0.10).	529 ERCPs, selective cannulation in 479 (90.5%). success rate after 100 ERCPs was 82%, 200 ERCPs 88%, 300 ERCPs 90% and 400 95%. Competence at n=79.	3
Verma[68] (2007)	Single center, retrospective, endoscopist without ERCP experience	1	Outcome: cannulation success of CBD in patients with native papillary anatomy. Learning curve per 50 procedures.	1097 ERCPs (697 as trainee, 400 as independent endoscopist). Success rates: baseline 42% and ≥80% after 350-400 procedures. Posttraining, the following 300 ERCPs >96% success.	4
Vitale[81] (2005)	*Single center, retrospective, fellows with different experience in ERCP (training period 6 to 14 months)*	13	Type of ERCP, success of cannulation of intended duct per 3-month period, faculty success where fellows failed	2008 ERCPs performed by fellows (median 135, range 53-351). first period, 95% diagnostic ERCPs, last period 95% therapeutic. Cannulation success improved from 77.3% in first period to 84.4% in the 4th period. After 7.1 months and 102 ERCPs, cannulation of >85%.	4

Table 5C. *Learning curves and threshold numbers as a measure of performance in EUS.*

Study	Setting and participants	n participants	Assessment method – outcome measures	Results	LOE
Wani[70] (2013)	Multicenter, prospective, trainees without prior EUS experience	5	Outcome: performance per anatomic station. Learning curves through cumulative sum (cusum) analysis. Acceptable and unacceptable failure rates: 0.10 and 0.20.	n = 1412 EUS examinations (median per trainee 295, range 175-402). Acceptable overall performance based on scores per anatomic station for 2 trainees after EUS 255 and 295. 2 trainees trend to competence (n=255 and 196), 1 needed ongoing observation after n=402.	2b
Meenan[72] (2003)	*Single center, prospective, GI trainees and nurse endoscopist with different experience in EUS*	4 & 1	Assessment of the ability to reproduce set views from esophagus, stomach and duodenum, points per structure (18, 8 and 11, respectively), 5 examinations each	Previous EUS experience: observed ranged from 55 to 170 procedures, performed from 25 to 124. Mean overall score per trainee: 37.4; 32; 29.7; 23.3. Nurse: 12.5 out of 18 (esophagus alone).Correlation between n performed and score for duodenal views (p<0.01). Competence after 25 procs in esophagus, 35 in the stomach, 78 in duodenum.	3

threshold numbers needed to gain competence

Nine studies reported numbers needed to gain competence in different procedures in gastrointestinal endoscopy. [72-81] These studies are shown in Table 5 as well. Two studies handled both EGD and colonoscopy [73, 80], whereas most of the studies pertained to colonoscopy alone. [74, 75, 77-79] There were two single studies on sigmoidoscopy and ERCP. [76, 81] The level of evidence was moderate for most studies due to the designs and numbers of procedures evaluated. Only three groups performed studies (regarding EGD and colonoscopy) with a prospective design and a considerable amount of trainees evaluated, resulting in LOE 2. [73, 74, 80] These will be discussed in further detail.

For EGD, competence was measured in two ways: intubation of the oesophagus and reaching a sufficient score on the Global Assessment of Gastrointestinal Endoscopic Skills (GAGES). One group demonstrated an 80% success rate of oesophageal intubation after 100 procedures, whereas another study concluded a plateau in the GAGES score after 50 procedures. [73, 80] Concerning colonoscopy, competence was measured trough CIR and scores on the GAGES form as well. Two studies concluded that 100 colonoscopies was insufficient for reaching a >90% CIR [73, 74], whereas the GAGES score displayed a plateau score at n=75 procedures. [80] All studies confirmed that performance of trainees increased with experience.

assessment and grading of performance

Nine studies described the development and evaluation of assessment tools for colonoscopy (n=6), sigmoidoscopy (n=1), both (n=1) and both colonoscopy and EGD (n=1). [82-90] These are shown in Table 6. All evidence level 2 studies focused on colonoscopy, flexible sigmoidoscopy or both, had a prospective design, and reported on 18 to 162 participants. [82, 86-89] The British Direct Observation of Procedural Skills (DOPS) appears effective for evaluation of competence for already registered endoscopists. [82] The Mayo Colonoscopy Skills Assessment Tool (MCSAT) was more effective in discriminating different experience levels, and therefore applicable in training settings. [87] Two studies reported on some sort of video assessment of endoscopic skills. [88, 89] The tri-split video recording assessment tool proved to be valid, but reliability was lacking. [88] The other study on video assessment described the development of an assessment tool for sigmoidoscopy withdrawals in a series of five experiments. [89] They concluded that the sequential assessment of five withdrawals led to the highest agreement. However, all procedures included in this video study were performed by experienced endoscopists. Some assessment tools were applicable in training situations, while others were only evaluated in a setting with experienced endoscopists. This difference makes it therefore difficult to compare the assessment tools.

Table 6. *Development and evaluation of assessment tools for colonoscopy, sigmoidoscopy and EGD.*

Study	Setting and participants	n participants	Assessment method – outcome measures	Results	LOE
Barton[82] (2012)	Candidates for bowel cancer screening program, prospective	162	Evaluation of colonoscopy by experts through DOPS. Grading performance through DOPS over 2 consecutive colonoscopies, global expert assessment, reliability and validity	193 assessments, 46 excluded, 147 candidates remained. Reliability of 0.81 (high). Validity: 72.6% of candidates and 92.9% of assessors experienced DOPS as valid. Global expert assessment and DOPS grade were similar in 97%. Variance with candidates.	2b
Shah[88] (2002)	Prospective, single center, endoscopists with different endoscopy experience	18	Tri-split video of colonoscopy (endoscopists hands, endoscopy view, magnetic endoscopic imaging) during insertion. Scoring system for instrument controls, insertion tube, depth of insertion	22 colonoscopies scored by each of 3 observers. 4 endoscopists with <100 colonoscopies experience, 3 with 250, 2 with 500 and 9 with >1000 experience. Significant differences between scorers (p<0.001). Good validity (differences for competence categories, p<0.001). Good interobserver agreement and correlation between individual scores and global assessment of competence (p<0.001). Reliability is lacking.	2b
Sedlack[87] (2010)	Single center, prospective, trainees	41	Development and evaluation of colonoscopy assessment tool, assessment of colonoscopy skills, validity testing, correlation between item scores and overall	3936 MCSAT completed (62%) by 58 staff members. Correlation between average and overall cognitive and motor scores (0.79 and 0.88, respectively, p<0.01). Difference in scores related to experience (p<0.01)	2b

Table 6. *Continued*

Study	Setting and participants	n participants	Assessment method – outcome measures	Results	LOE
Sarker[86] (2008)	Multicenter, prospective, trainees and consultants	21	Assessment of skills by Likert scale, hierarchical task analysis. Assessment of flexible sigmoidoscopy and colonoscopy, generic technical skills, two assessors, validity and reliability	135 endoscopies (75 fles sig, 60 colonoscopies) assessed, 9 consultants and 12 trainees. Cronbach's alpha for flex sig (generic vs. specific): 0.81 and 0.79 (p<0.05). For colonoscopy generic vs. specific 0.85 and 0.80 (p<0.05). Construct validity for experience in flex sig and colonoscopy: generic vs. specific p=0.005 and 0.003 and p=0.012 and 0.004.	2b
Thomas-Gibson[89] (2006)	UK flexible sigmoidoscopy screening trial, experienced endoscopists	43	Development of scoring system for assessment of accuracy for flexible sigmoidoscopy, series of 5 experiments for video scoring of extubation, up to 6 scorers per experiment. First two experiments scores on VAS, last three on 5-point Likert scale. Different items of extubation assessed	Overall ICC for first 2 experiments: 0.10. In experiment 3, all scores varied significantly between scorers (p<0.05). In experiment 4, overall parameters scored first with higher ICC (quality improved from 0.45 to 0.72). Experiment 5: sequential assessment of videos vs. random. Agreement was low for individual extubations (ICC 0.13) but high for series of 5 extubations (ICC 0.89).	2b
Boyle[83] (2012)	Multicenter, prospective, trainees and consultants	27	Development and evaluation of colonoscopy assessment tool. Assessment of elective colonoscopies, dichotomous checklist and global assessment	81 procedures (24 by 8 consultant and 57 by 19 trainees) assessed. Significant differences in overall score for novice, intermediate and expert endoscopists (p<0.001). Checklist: 3 items distinguished between two or more groups, global assessment: 6 items distinguished between three levels	3
Vassiliou[90] (2010)	Multi center, prospective, novice and experienced endoscopists	139	Development and evaluation of GAGES for EGD and colonoscopy, reliability and validity. Assessment by operator, attending and observer. Novices and experienced endoscopists	Data on attending and observer evaluations: 18 EGD and 13 colonoscopies. Data on attending and participants: 77 EGD and 57 colonoscopies. ICC for attendings and observers for both tools were 0.96 and 0.97. ICC for attending and participants: 0.78 and 0.89. Validity for both tools (p<0.001).	3
Hope[84] (2013)	Single center, prospective, surgical residents	100	Assessment of colonoscopy performance by two assessments tools (8 objective criteria vs 10 generic and specific endoscopic skills). Performance per PGY level	100 colonoscopies performed by residents, 89 assessed (72% PGY-3). Tool 1 showed on some items significant differences for PGY-level (p<0.05) but not for all. Tool 2 as well (p<0.05). Both tools not for all categories significantly different; they show improvement with experience but not for all items.	3

Table 6. *Continued*

Study	Setting and participants	n participants	Assessment method – outcome measures	Results	LOE
Mohamed[85] (2011)	Single center, prospective, first year GI fellows and third year surgery fellows	7	Assessment of DOPS reliability for trainees for colonoscopy. Comparison of 7-step model for training (GI fellows) colonoscopy with master-apprentice model (surgical fellows)	4 GI fellows with experience of 30 colonoscopies, 3 surgery fellows with 5 colonoscopies experience. Scores for GI fellows and surgical were comparable. Test-retest of DOPS wide range of correlations, no reliability.	4

training models

Finally, seven studies reported on different kinds of training models for colonoscopy (n=4), sigmoidoscopy (n=1) and EGD (n=2). [91-97] Table 7 provides an overview of these studies. Two groups described the evaluation of the accelerated colonoscopy training course (ACTC) as it is carried out in the UK. [96, 97] Both concluded that performance in knowledge, colonoscopy performance, and Direct Observation of Procedural Skills (DOPS) scores improved significantly after the training week. Thomas-Gibson et al. added an evaluation at a median follow-up of 9 months. There were however no differences between post-training assessment and follow-up. A different training model was the 'gastroenterological education – training endoscopy' (GATE) model. [91] This training model showed improvement in post-test results and simulator performance. A German group tried to identify predictors for performance in a 1-week training course by psychological and psychomotor tests. [94] The training week resulted in improved performance, but only one specific (double labyrinth) test was identified as a predictor for improvement in performance.

Table 7. *Training models for colonoscopy, sigmoidoscopy and EGD.*

Study	Setting and participants	n participants	Assessment method – outcome measures	Results	LOE
Götzberger[91] (2011)	*Multicenter, training courses, trainees with different levels of endoscopy experience*	98	Evaluation of GATE model (basic and therapeutic course), pre- and postcourse knowledge, acceptance, simulator assessment, colonoscopy	78 trainees provided complete data sets. Acceptance for both courses: 88% would advise participation. 77% underlined realism. Improvement in pre-and post-test results (p<0.001). Simulator assessment: reduction in time needed for procedure (p<0.01).	2b
Harewood[92] (2008)	*Single center, RCT, trainees with experience of 500-600 colonoscopies*	4	Evaluation of effect of systematic feedback on performance (1-on-1, monthly) in colonoscopy. Evaluation of CIR and PDR.	581 colonoscopies performed, 296 in feedback group, 285 in control group. After feedback, 364 procs performed (211 feedback, 153 control). In feedback group, CIR improved from 72.9 to 83.4% (p=0.04). In control group CIR from 78 to 71.9% (p=0.2). PDR not significantly different (p=0.2 vs p=0.5).	2b

Table 7. *Continued*

Study	Setting and participants	n participants	Assessment method – outcome measures	Results	LOE
Neumann[94] (2005)	*Multicenter, prospective, trainees without prior endoscopic experience*	58	Evaluation of psychological, psychomotor, cognitive tests and subjective assessment of trainer before and after 1-week training course in EGD	58 trainees from 12 centers evaluated. All showed significant improvement in performance ($p<0.001$). Only double labyrinth test was associated with improvement (OR 11.5, $p=0.035$), expert assessment of at least moderate improvement (OR 41.5, $p=0.018$)	2b
Suzuki[96] (2006)	*Multicenter, prospective, trainees with different endoscopic experience*	50	Assessment of accelerated colonoscopy training course: multiple choice questions (MCQ), simulator and hands-on training, evaluation DOPS	50 trainees attended. MCQ score improved from 57 to 66% posttraining ($p<0.001$). Median of 15 live colonoscopies performed: DOPS showed improvement on all aspects ($p=0.007$). Global score improved from 1.7 to 2.0 ($p<0.001$).	2b
Thomas-Gibson[97] (2007)	*Multicenter, prospective, trainees with different endoscopic experience*	21	Assessment of accelerated colonoscopy training course, MCQ, simulator sessions, hands on training, live case assessment, trisplit video assessment. Follow up: MCQ, simulator, live case assessment, trisplit video assessment	13 trainees with <200 colonoscopies, 5 with 200-500 colonoscopies, 3 with 500-1000 colonoscopies; 16 follow-up. Significant improvement in MCQ scores ($p<0.001$), simulator test times (diagnostic vs therapeutic $p=0.02$ vs $p=0.003$). DOPS scores improved significantly on all aspects expect general approach. Trisplit video: 1 scorer gave higher grades ($p=0.008$), 1 did not ($p=0.11$). No difference between post-training week scores and at follow up, except simulator performance.	2b
Neumann[93] (2003)	*Multi center, prospective, trainees with different endoscopic experience.*	56	Evaluation of training course for EGD with Erlangen Endo-Trainer, endoscopy Score Cards, tests on each day, self-assessment, video assessments	Assessment of day 1 and 5 training significantly different ($p<0.001$) for both specialist and self-assessment. During entire training, differences in scores of self-assessment and expert assessment on 24 items ($p<0.05$).	3
Proctor[95] (1998)	*Single center, prospective evaluation, trainees without previous endoscopic experience*	10	Evaluation of teaching model for flexible sigmoidoscopy, 8 components with two levels of competence and 1 level non-competent. Comparison with independent observer method	120 of 128 sigmoidoscopies evaluated using teaching model method; 73 (60.8%) competent, 47 incompetently performed. 1 component excluded due to lack of variation; 6 of 7 remaining components were associated with competence in teaching model. 50 of 120 procedures Overall Competence Score; correlation between this and teaching model was 0.71 ($p<0.001$).	3

One RCT evaluated the impact of systematic feedback on colonoscopy performance. [92] Although only four trainees were evaluated, there was a significant improvement in CIR performance in the feedback group, while the control group showed no improvement.

DISCUSSION

forward viewing flexible endoscopy procedures

Gastrointestinal endoscopic procedures are fairly complex. The sole use of the classic master-apprentice model for teaching endoscopy is nowadays less accepted. The use of simulators in the early training phase is gaining acceptance and several VR endoscopy simulators have been validated (Table 2). The GI Mentor, AccuTouch, and Endo TS-1 were shown to have good validity. [4, 8-10, 12-16, 19-22, 25, 27, 28] These can thus be considered as realistic devices that have discriminative abilities for distinguishing dexterity and competence levels in flexible endoscopy. Based on these LOE 2b studies, the grade of recommendation for the validity of the mentioned mechanical simulators is B.

Following validation, the impact of simulator training on learning curves needs to be assessed. A VR simulator with good validity, but not improving performance after repeated exercise, is not suitable for implementing in a training program. Three studies with LOE 1b provided evidence for the positive effect of simulator training in novices in flexible endoscopy, measured in terms of both virtual reality as well as live endoscopy. [37, 42, 53] Two of these were well-designed randomized multicenter trials comparing the combination of simulator- and bedside-training versus bedside training alone for colonoscopy training of novices. These studies demonstrated that simulator training is effective. [37, 42] Several studies on simulator learning curves for EGD, sigmoidoscopy and colonoscopy gained a LOE of 2a or 2b. [31, 32, 34, 36, 38-41, 43-45, 47, 49-52] Based on this evidence, one can conclude that simulator training is complementary to patient-based learning and is useful in the early training phase, resulting in a grade of recommendation B.

The four studies that reported on the use of a simulator as a competence assessment tool showed diverging results. [54-57] Therefore, no grade of recommendation is given.

Elaborating further on the learning curve, the next step is (continuous) assessment of a trainee's performance during patient-based training. The currently available recommendations and guidelines focus mainly on minimum numbers as a threshold for competence [73-80]. However, outcomes and proposed minimum numbers for flexible endoscopic procedures vary widely. Nowadays there is a tendency to define more objective criteria for competence. Two large prospective single-centre studies with LOE 1b provided evidence for the use of an assessment form as a measure of competence, respectively the Mayo Colonoscopy Skills Assessment Tool (MCSAT) and the Rotterdam Assessment Form for Colonoscopy (RAF-C). [61, 65] The learning curves obtained in these studies were similar. Both forms are good methods to continuously assess performance, resulting in a grade of recommendation A. The DOPS on the other hand is more appropriate for assessment of 'end-stage' competence. [82]

Overall, some high-quality studies have been performed for each individual step in training, providing valuable information on the effect of simulator training, learning curves and assessment methods. The most and best evidence for all these stages regarding basic flexible endoscopy is available for colonoscopy. However, one can imagine that some results can be extrapolated to other basic GI endoscopy procedures as well, since the techniques are comparable.

ERCP

One of the most challenging procedures with high complication rates in GI endoscopy is ERCP. It takes a great deal of training and a large number of procedures to reach competence. However, little is known about the learning curve for trainees in ERCP. A number of questions remain unanswered when it comes to the shape of the learning curve, the number of procedures needed to gain competence, and the definition of competence itself. The six studies on learning curves in ERCP varied widely in design, number of trainees and procedures included, as well as outcome, resulting in a large heterogeneity among them. Successful cannulation in >85% of the patients was seen after a number of 100 to 185 ERCPs. Due to the heterogeneity, a grade of recommendation is not provided. There would be a great benefit if part of the learning curve for endoscopists could be accomplished by training on simulator models. In reality the number of available simulators for training in ERCP is limited. Seven validation studies have been performed in an attempt to validate 6 different ERCP simulator models. A 2b level of evidence was reached in five studies; two studies were merely feasibility studies. The GI Mentor is the only validated VR simulator for ERCP. [7, 30] The face and construct validity was demonstrated in these two studies and although it received lower scores than the ex vivo or live porcine model in a head-to-head comparison, it was considered the easiest of all ERCP simulator models to incorporate in a training curriculum. [7, 30] The live porcine model was validated only once in comparison to the ex vivo model and the GI Mentor in the same study. [30] The ex vivo simulators and purely mechanical simulators are highly comparable among each other and achieve similar results. All of these models require a real endoscope to be introduced to reach a papilla which is either a synthetic or an ex vivo papilla located in a mechanical tube representing the duodenum or an ex vivo duodenum. Overall these ERCP simulator models receive the highest scores on realism. In total four studies was performed reaching a 2b level of evidence with a fairly good concordance resulting in a grade B recommendation. Only one learning curve study was performed, demonstrating higher success in the simulator group. [46] Since this was the only learning curve study on simulator-based ERCP in the literature, no grade of recommendation can be provided. There were no studies found on validated competence assessment tools to objectify performance in ERCP. The most common performance parameter is cannulation success rate. This only partly reflects the extent, therapeutic intent and diversity of a therapeutic procedure like an ERCP.

endosonography

Endosonography or endoscopic ultrasound (EUS) is widely practiced with an increasing number of therapeutic possibilities since the first reports of transgastric drainage of pseudocysts by Grimm et al. [98] This makes EUS more complex. Especially the therapeutic procedures have a marked overlap with ERCP and demand a great deal of experience. There are only a few reports on simulator-based training in EUS. [5, 6] Training diagnostic and interventional EUS seems logical and feasible in a live porcine model but no formal attempt at validation has been made. No grade of recommendation can be given based on these studies. A learning effect by repeated exercise and improvement of performance during EUS procedures in the live porcine model itself was documented in one study. [35] There is a lack of scientific evidence of transfer of competence to a patient-based setting. There is an even greater scarcity of evidence on learning curves and numbers to reach competence in EUS. Two studies were performed that both included five trainees. The first study included only radial EUS. [72] They reported no additional effect of observing large numbers of procedures; the largest benefit was achieved during hands-on training. There is only one LOE 2b study performed. [70] The learning curves differed considerably among the five trainees. These studies demonstrated the substantial need for much more training than the 150 procedures recommended by the American Society for Gastrointestinal Endoscopy (ASGE) in order to reach proficiency. It leads to a grade C recommendation.

limitations

The heterogeneity of the studies regarding forward viewing endoscopy limits the conclusions that can be drawn. This systematic review covers a broad range of studies regarding training and assessment in GI endoscopy. This broad approach automatically results in a large variety of methodology, devices used and endpoints measured. This hampers head-to-head comparison of individual studies. Another limitation concerns the fact that all studies focused on specific aspects of the endoscopic procedure, instead of on overall performance, which is both overall competence assessment from novice to experienced, certified endoscopy, as well as expert levels for specific procedures.

The evidence in the literature on learning curves and competence measures for ERCP is highly heterogeneous. This makes it impossible to provide a level of recommendation. Also, cannulation success rates do relate to improved performance but do not entirely reflect the diversity of a complex procedure like ERCP. No solid data are currently available on other aspects of therapeutic interventions related to learning curves and benchmarks in ERCP. As of yet, no validated competence assessment tools have been developed for ERCP. This should be a prerequisite before attempts to define learning goals and benchmarks are made.

future research

Future research, based on the presented evidence in this review, should therefore include a complete training program. We propose a pre-patient curriculum using simulator training. The transfer of simulation skills to patient-based procedures needs to be further explored. Simulation training needs to be followed by continuous assessment of patient-based endoscopies to provide individual and group learning curves and after a period of time, (repeated) overall assessments of performance by an expert. Therefore, the development of validated assessment tools is necessary and the effect of expert assessments on daily practice needs to be measured.

With respect to ERCP, there is a rationale to start training using simulators. There is however no evidence yet as to what extent or performance level simulator-based training has to be carried out. The next step would be to investigate the transfer of skills to patient-based training. These research objectives seem to be clear goals for future research. There is a need for the development of validated objective assessment tools in ERCP to document progress in training and finally proficiency. Benchmarks can be set using the same assessment tools in ERCP performed by experts.

The evidence on training and competence assessment in EUS is extremely scarce. Although training in a live porcine model seems logical, in the current era of evidence-based medicine, validation studies should be carried out to establish the degree of realism and training potential. Current threshold numbers for training appear to be inadequate, but the available data are sparse. We seem to be far away from establishing benchmarks for competence in EUS and validated assessment tools are lacking.

general conclusions and recommendations

Based on the presented evidence, we propose implementation of simulator training in gastrointestinal endoscopic training curricula. Regarding basic flexible endoscopy (EGD, sigmoidoscopy and colonoscopy), simulator-based training has proven its value and it is justifiable to start a pre-patient training course using a validated simulator. The extent to which simulator-based training should be carried out is still a matter of debate. Furthermore, objective outcome parameters should be measured continuously in patient-based training. This provides insight in the learning curve and is preferable to threshold numbers. The MCSAT, RAF-C and DOPS assessment forms seem to be the best forms to document progress or proficiency levels. Regarding ERCP training, we would recommend a pre-patient training curriculum using a validated simulator as well. Evidence for evaluation of learning curves and continuous assessment in ERCP is scarce. This makes competency based training difficult. The available data support prolonged training, at least to a larger extent than current upheld threshold numbers in most countries. The results so far may hopefully stimulate further research. The evidence on endosonography training and competence is yet the least investigated. A pre-patient training curriculum is logical and attractive. However, the evidence is too scarce to give recommendations at this moment.

REFERENCES

1. Grantcharov TP, Reznick RK. Teaching procedural skills. BMJ 2008;336:1129-1131
2. Centre for EBM Levels of Evidence. Oxford Centre for Evidence-based Medicine - levels of evidence. In: http://www.cebm.net/index.aspx?o=1025; 2011
3. Carter FJ, Schijven MP, Aggarwal R, et al. Consensus guidelines for validation of virtual reality surgical simulators. Surg Endosc 2005;19:1523-1532
4. Datta V, Mandalia M, Mackay S, Darzi A. The PreOp flexible sigmoidoscopy trainer: Validation and early evaluation of a virtual reality based system. Surg Endosc 2002;16:1459-1463
5. Bhutani MS, Aveyard M, Stills Jr HF. Improved model for teaching interventional EUS. Gastrointest Endosc 2000;52:400-403
6. Bhutani MS, Hoffman BJ, Hawes RH. A swine model for teaching endoscopic ultrasound (EUS) imaging and intervention under EUS guidance. Endoscopy 1998;30:605-609
7. Bittner Iv JG, Mellinger JD, Imam T, Schade RR, MacFadyen Jr BV. Face and construct validity of a computer-based virtual reality simulator for ERCP. Gastrointest Endosc 2010;71:357-364
8. Fayez R, Feldman LS, Kaneva P, Fried GM. Testing the construct validity of the Simbionix GI Mentor II virtual reality colonoscopy simulator metrics: module matters. Surgical endoscopy 2010;24: 1060-1065
9. Felsher JJ, Olesevich M, Farres H, et al. Validation of a flexible endoscopy simulator. Am J Surg 2005; 189:497-500
10. Ferlitsch A, Glauninger P, Gupper A, et al. Evaluation of a virtual endoscopy simulator for training in gastrointestinal endoscopy. Endoscopy 2002;34:698-702
11. Frimberger E, von Delius S, Rosch T, et al. A novel and practicable ERCP training system with simulated fluoroscopy. Endoscopy 2008;40:517-520
12. Grantcharov TP, Carstensen L, Schulze S. Objective assessment of gastrointestinal endoscopy skills using a virtual reality simulator. JSLS 2005;9:130-133
13. Haycock AV, Bassett P, Bladen J, Thomas-Gibson S. Validation of the second-generation Olympus colonoscopy simulator for skills assessment. Endoscopy 2009;41:952-958
14. Kim S, Spencer G, Makar GA, et al. Lack of a discriminatory function for endoscopy skills on a computer-based simulator. Surg Endosc Interv Tech 2010;24:3008-3015
15. Koch AD, Buzink SN, Heemskerk J, et al. Expert and construct validity of the Simbionix GI Mentor II endoscopy simulator for colonoscopy. Surg Endosc 2008;22:158-162
16. Koch AD, Haringsma J, Schoon EJ, de Man RA, Kuipers EJ. A second-generation virtual reality simulator for colonoscopy: Validation and initial experience. Endoscopy 2008;40:735-738
17. Leung JW, Lee JG, Rojany M, Wilson R, Leung FW. Development of a novel ERCP mechanical simulator. Gastrointest Endosc 2007;65:1056-1062
18. Leung JW, Wang D, Hu B, Lim B, Leung FW. A head-to-head hands-on comparison of ERCP mechanical simulator (EMS) and ex-vivo porcine stomach model (PSM). J Intervent Gastroenterol 2011;1:108-113
19. MacDonald J, Ketchum J, Williams RG, Rogers LQ. A lay person versus a trained endoscopist: can the preop endoscopy simulator detect a difference? Surg Endosc 2003;17:896-898
20. Mahmood T, Darzi A. A study to validate the colonoscopy simulator: It is usefully discriminatory for more than one measurable outcome. Surg Endosc 2003;17:1583-1589
21. McConnell RA, Kim S, Ahmad NA, et al. Poor discriminatory function for endoscopic skills on a computer-based simulator. Gastrointest Endosc 2012
22. Moorthy K, Munz Y, Jiwanji M, et al. Validity and reliability of a virtual reality upper gastrointestinal simulator and cross validation using structured assessment of individual performance with video playback. Surg Endosc 2004;18:328-333
23. Neumann M, Mayer G, Ell C, et al. The Erlangen endo-trainer: Life-like simulation for diagnostic and interventional endoscopic retrograde cholangiography. Endoscopy 2000;32:906-910

24. Plooy AM, Hill A, Horswill MS, et al. Construct validation of a physical model colonoscopy simulator. Gastrointest Endosc 2012;76:144-150
25. Ritter EM, McClusky Iii DA, Lederman AB, Gallagher AG, Smith CD. Objective psychomotor skills assessment of experienced and novice flexible endoscopists with a virtual reality simulator. J Gastrointest Surg 2003;7:871-878
26. Sedlack RE, Baron TH, Downing SM, Schwartz AJ. Validation of a colonoscopy simulation model for skills assessment. Am J Gastroenterol 2007;102:64-74
27. Sedlack RE, Kolars JC. Validation of a computer-based colonoscopy simulator. Gastrointest Endosc 2003;57:214-218
28. Sugden C, Aggarwal R, Banerjee A, et al. The development of a virtual reality training curriculum for colonoscopy. Ann Surg 2012;256:188-192
29. von Delius S, Thies P, Meining A, et al. Validation of the X-Vision ERCP Training System and Technical Challenges During Early Training of Sphincterotomy. Clin Gastroenterol Hepatol 2009;7:389-396
30. Sedlack R, Petersen B, Binmoeller K, Kolars J. A direct comparison of ERCP teaching models. Gastrointest Endosc 2003;57:886-890
31. Tuggy ML. Virtual reality flexible sigmoidoscopy simulator training: impact on resident performance. J Am Board Fam Pract 1998;11:426-433
32. Van Sickle KR, Buck L, Willis R, et al. A multicenter, simulation-based skills training collaborative using shared GI mentor II systems: Results from the Texas association of surgical skills laboratories (TASSL) flexible endoscopy curriculum. Surg Endosc Interv Tech 2011;25:2980-2986
33. Ahad S, Boehler M, Schwind CJ, Hassan I. The effect of model fidelity on colonoscopic skills acquisition. A randomized controlled study. J Surg Educ 2013;70:522-527
34. Ahlberg G, Hultcrantz R, Jaramillo E, Lindblom A, Arvidsson D. Virtual reality colonoscopy simulation: A compulsory practice for the future colonoscopist? Endoscopy 2005;37:1198-1204
35. Barthet M, Gasmil M, Boustiere C, et al. EUS training in a live pig model: Does it improve echo endoscope hands-on and trainee competence? Endoscopy 2007;39:535-539
36. Buzink SN, Koch AD, Heemskerk J, et al. Acquiring basic endoscopy skills by training on the GI Mentor II. Surg Endosc Interv Tech 2007;21:1996-2003
37. Cohen J, Cohen SA, Vora KC, et al. Multicenter, randomized, controlled trial of virtual-reality simulator training in acquisition of competency in colonoscopy. Gastrointest Endosc 2006;64:361-368
38. Eversbusch A, Grantcharov TP. Learning curves and impact of psychomotor training on performance in simulated colonoscopy: A randomized trial using a virtual reality endoscopy trainer. Surg Endosc 2004;18:1514-1518
39. Ferlitsch A, Schoefl R, Puespoek A, et al. Effect of virtual endoscopy simulator training on performance of upper gastrointestinal endoscopy in patients: A randomized controlled trial. Endoscopy 2010;42:1049-1056
40. Gerson LB, Van Dam J. A prospective randomized trial comparing a virtual reality simulator to bedside teaching for training in sigmoidoscopy. Endoscopy 2003;35:569-575
41. Giulio ED, Fregonese D, Casetti T, et al. Training with a computer-based simulator achieves basic manual skills required for upper endoscopy: A randomized controlled trial. Gastrointest Endosc 2004;60:196-200
42. 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
43. Kaltenbach T, Leung C, Wu K, et al. Use of the colonoscope training model with the colonoscope 3d imaging probe improved trainee colonoscopy performance: A pilot study. Dig Dis Sci 2011;56:1496-1502

44. Kruglikova I, Grantcharov TP, Drewes AM, Funch-Jensen P. The impact of constructive feedback on training in gastrointestinal endoscopy using high-fidelity virtual-reality simulation: A randomised controlled trial. Gut 2010;59:181-185
45. Kruglikova I, Grantcharov TP, Drewes AM, Funch-Jensen P. Assessment of early learning curves among nurses and physicians using a high-fidelity virtual-reality colonoscopy simulator. Surg Endosc Interv Tech 2010;24:366-370
46. Lim BS, Leung JW, Lee J, et al. Effect of ERCP mechanical simulator (EMS) practice on trainees ERCP performance in the early learning period: US multicenter randomized controlled trial. Am J Gastroenterol 2011;106:300-306
47. Mahmood T, Darzi A. The learning curve for a colonoscopy simulator in the absence of any feedback: No feedback, no learning. Surg Endosc 2004;18:1224-1230
48. Maiss J, Millermann L, Heinemann K, et al. The compactEASIE(registered trademark) is a feasible training model for endoscopic novices: A prospective randomised trial. Dig Liver Dis 2007;39:70-78
49. Sedlack RE, Kolars JC. Computer simulator training enhances the competency of gastroenterology fellows at colonoscopy: Results of a pilot study. Am J Gastroenterol 2004;99:33-37
50. Sedlack RE, Kolars JC, Alexander JA. Computer simulation training enhances patient comfort during endoscopy. Clin Gastroenterol Hepatol 2004;2:348-352
51. Snyder CW, Vandromme MJ, Tyra SL, Hawn MT. Retention of colonoscopy skills after virtual reality simulator training by independent and proctored methods. Am Surg 2010;76:743-746
52. Park J, MacRae H, Musselman LJ, et al. Randomized controlled trial of virtual reality simulator training: transfer to live patients. Am J Surg 2007;194:205-211
53. Ende A, Zopf Y, Konturek P, et al. Strategies for training in diagnostic upper endoscopy: A prospective, randomized trial. Gastrointest Endosc 2012;75:254-260
54. Elvevi A, Cantu P, Maconi G, Conte D, Penagini R. Evaluation of hands-on training in colonoscopy: Is a computer-based simulator useful? Dig Liver Dis 2012;44:580-584
55. Moorthy K, Munz Y, Orchard TR, et al. An innovative method for the assessment of skills in lower gastrointestinal endoscopy. Surg Endosc 2004;18:1613-1619
56. Phitayakorn R, Marks JM, Reynolds HL, Delaney CP. Expert benchmark for the GI Mentor II. Surg Endosc Interv Tech 2009;23:611-614
57. Sarker SK, Albrani T, Zaman A, Kumar I. Procedural performance in gastrointestinal endoscopy: live and simulated. World J Surg 2010;34:1764-1770
58. Biau DJ, Williams SM, Schlup MM, Nizard RS, Porcher R. Quantitative and individualized assessment of the learning curve using LC-CUSUM Review. Br J Surg 2008;95:925-929
59. Chung JI, 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
60. Jowell PS, Baillie J, Branch MS, et al. Quantitative assessment of procedural competence: A prospective study of training in endoscopic retrograde cholangiopancreatography. Ann Intern Med 1996;125:983-989
61. Koch AD, Haringsma J, Schoon EJ, De Man RA, Kuipers EJ. Competence measurement during colonoscopy training: The use of self-assessment of performance measures. Am J Gastroenterol 2012;107:971-975
62. Lee SH, Chung IK, Kim SJ, et al. An adequate level of training for technical competence in screening and diagnostic colonoscopy: a prospective multicenter evaluation of the learning curve. Gastrointest Endosc 2008;67:683-689
63. Marshall JB. Technical proficiency of trainees performing colonoscopy: A learning curve. Gastrointest Endosc 1995;42:287-291
64. Parry BR, Williams SM. Competency and the colonoscopist: a learning curve. AUST NEW ZEALAND J SURG 1991;61:419-422

65. Sedlack RE. Training to competency in colonoscopy: Assessing and defining competency standards. Gastrointest Endosc 2011;74:355-366+366.e351-366.e352
66. Selvasekar CR, Holubar SD, Pendlimari R, et al. Assessment of screening colonoscopy competency in colon and rectal surgery fellows: A single institution experience. J Surg Res 2012;174:e17-e23
67. Tassios PS, Ladas SD, Grammenos I, Demertzis K, Raptis SA. Acquisition of competence in colonoscopy: The learning curve of trainees. Endoscopy 1999;31:702-706
68. Verma D, Gostout CJ, Petersen BT, et al. Establishing a true assessment of endoscopic competence in ERCP during training and beyond: a single-operator learning curve for deep biliary cannulation in patients with native papillary anatomy. Gastrointest Endosc 2007;65:394-400
69. Waller HM, Connor SJ. Cumulative sum (Cusum) analysis provides an objective measure of competency during training in endoscopic retrograde cholangio-pancreatography (ERCP). Hpb 2009;11:565-569
70. Wani S, Cote GA, Keswani R, et al. Learning curves for EUS by using cumulative sum analysis: Implications for American Society for Gastrointestinal Endoscopy recommendations for training. Gastrointest Endosc 2013;77:558-565
71. Watkins JL, Etzkorn KP, Wiley TE, DeGuzman L, Harig JM. Assessment of technical competence during ERCP training. Gastrointest Endosc 1996;44:411-415
72. Meenan J, Anderson S, Tsang S, et al. Training in radial EUS: What is the best approach and is there a role for the nurse endoscopist? Endoscopy 2003;35:1020-1023
73. Cass OW, Freeman ML, Peine CJ, Zera RT, Onstad GR. Objective evaluation of endoscopy skills during training. Ann Intern Med 1993;118:40-44
74. Chak A, Cooper GS, Blades EW, Canto M, Sivak M.V, Jr. Prospective assessment of colonoscopic intubation skills in trainees. Gastrointest Endosc 1996;44:54-57
75. Church J, Oakley J, Milsom J, Strong S, Hull T. Colonoscopy training: The need for patience (patients). ANZ J Surg 2002;72:89-91
76. Hawes R, Lehman GA, Hast J. Training resident physicians in fiberoptic sigmoidoscopy. How many supervised examinations are required to achieve competence? Am J Med 1986;80:465-470
77. Leyden JE, Doherty GA, Hanley A, et al. Quality of colonoscopy performance among gastroenterology and surgical trainees: A need for common training standards for all trainees? Endoscopy 2011;43:935-940
78. Spier BJ, Benson M, Pfau PR, et al. Colonoscopy training in gastroenterology fellowships: determining competence. Gastrointest Endosc 2010;71:319-324
79. Spier BJ, Durkin ET, Walker AJ, et al. Surgical resident's training in colonoscopy: Numbers, competency, and perceptions. Surg Endosc Interv Tech 2010;24:2556-2561
80. Vassiliou MC, Kaneva PA, Poulose BK, et al. How should we establish the clinical case numbers required to achieve proficiency in flexible endoscopy? Am J Surg 2010;199:121-125
81. Vitale GC, Zavaleta CM, Vitale DS, et al. Training surgeons in endoscopic retrograde cholangiopancreatography. Surg Endosc Interv Tech 2006;20:149-152
82. 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
83. Boyle E, Al-Akash M, Patchett S, Traynor O, McNamara D. Towards continuous improvement of endoscopy standards: Validation of a colonoscopy assessment form. Colorectal Dis 2012;14:1126-1131
84. Hope WW, Hooks Iii WB, Kilbourne SN, et al. Assessing resident performance and training of colonoscopy in a general surgery training program. Surg Endosc Interv Tech 2013;27:1706-1710
85. Mohamed R, Shaheen AA, Raman M. Evaluation of colonoscopy skills - How well are we doing? Can J Gastroenterol 2011;25:198-200
86. Sarker SK, Albrani T, Zaman A, Patel B. Procedural performance in gastrointestinal endoscopy: An assessment and self-appraisal tool. Am J Surg 2008;196:450-455

87. 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.e1123
88. 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
89. Thomas-Gibson S, Rogers PA, Suzuki N, et al. Development of a video assessment scoring method to determine the accuracy of endoscopist performance at screening flexible sigmoidoscopy. Endoscopy 2006;38:218-225
90. Vassiliou MC, Kaneva PA, Poulose BK, et al. Global assessment of gastrointestinal endoscopic skills (GAGES): A valid measurement tool for technical skills in flexible endoscopy. Surg Endosc Interv Tech 2010;24:1834-1841
91. Gotzberger M, Rosch T, Geisenhof S, et al. Effectiveness of a novel endoscopy training concept. Endoscopy 2011;43:802-807
92. Harewood GC, Murray F, Winder S, Patchett S. Evaluation of formal feedback on endoscopic competence among trainees: The EFFECT trial. Ir J Med Sci 2008;177:253-256
93. Neumann M, Hahn C, Horbach T, et al. Score card endoscopy: A multicenter study to evalaute learning curves in 1-week courses using the Elangen Endo-Trainer. Endoscopy 2003;35:515-520
94. Neumann M, Meining A, Buerschaper C, et al. Training in GI endoscopy: Can we objectively predict the performance of trainees? A prospective correlation study. Z Gastroenterol 2005;43:445-450
95. Proctor DD, Price J, Dunn KA, et al. Prospective evaluation of a teaching model to determine competency in performing flexible sigmoidoscopies. Am J Gastroenterol 1998;93:1217-1221
96. Suzuki N, Thomas-Gibson S, Vance M, et al. Efficacy of an accelerated colonoscopy training week: Audit from one national colonoscopy training center in the UK. DIG ENDOSC 2006;18:288-293
97. Thomas-Gibson S, Bassett P, Suzuki N, et al. Intensive training over 5 days improves colonoscopy skills long-term. Endoscopy 2007;39:818-824
98. Grimm H, Binmoeller KF, Soehendra N. Endosonography-guided drainage of a pancreatic pseudocyst. Gastrointest Endosc 1992;38:170-171

Chapter 2

Aims and outline of the thesis

This thesis focuses on the development and measurement of skills during colonoscopy training of novice endoscopists and assessing competency levels for certification. Colonoscopy is a technique with significant burden on patients undergoing this procedure. Performing colonoscopy requires practical skill that is obtained by extensive training and has to be acquired over several hundred procedures. Trainees differ considerably in the rate at which they acquire the necessary skills, although evidence suggests that most will reach a basic competency level. This point is named procedural competence. Measures for colonoscopy competence are however ill defined. In most countries, threshold numbers for certification are used with hardly any regard for the quality of these procedures. For a colonoscopy with complete inspection of the mucosa, cecal intubation is a prerequisite. Cecal intubation rate is therefore often used as a measure of procedural competence. Generally, reaching the cecum in 90 out of 100 attempts is considered the minimum threshold for competence. However, the use of the cecal intubation rate as a primary endpoint provides no insight in the manner that the cecum was reached, in how much time, and with which effort and patient burden.

The aims of this thesis are to investigate the role of virtual reality (VR) endoscopy simulator training as a means to traverse the first and most burdensome part of the learning curve in flexible endoscopy, to explore the transfer of simulator-acquired skills to patient-based colonoscopy and to develop an assessment tool to gain insight in the learning curve of novice endoscopists and to create new benchmarks for colonoscopy training and competency measurement.

This thesis is divided in four sections and nine chapters. In section one, **Introduction and aims of this thesis**, the aims of this thesis are outlined. **Chapter 1** provides a systematic review on training and competence assessment in gastrointestinal (GI) endoscopy. We review the current evidence and expose knowledge gaps in GI endoscopy training, assessment and competency benchmarking.

In section two, **Simulator-based endoscopy – Validation and learning curves in colonoscopy**, the value of VR endoscopy simulator-based training is investigated. In **chapter 3** and **5**, two VR endoscopy simulators, the Simbionix GI Mentor II™ and the Olympus Endo TS-1™, are subjected to validation according to the principles of evidence-based science. *Face* and *construct validity* is investigated for both simulators being the most important validity pillars for simulators to be used in dexterity training. The first steps for novice endoscopists in flexible endoscopy are usually considered the most burdensome for both patients and trainees. It therefore seems most logical to train these first steps using simulators instead of real patients. In **chapter 4**, the effect of repeated training on the simulator is investigated. It is demonstrated that novices approach the performance of experienced endoscopists by training on the VR endoscopy simulators on both flexible endoscopy and dexterity tasks while experts do not improve after repeated exercises.

In section three, **Transfer of skills to patient-based endoscopy**, the procedural skills that have been acquired while training on the simulators are investigated during patient-based colonoscopy. It is of course of the utmost importance that skills acquired on the simulator do actually transfer to procedural skills that are useful and similar to the ones needed for flexible endoscopy. In **chapter 6** and **7**, it is demonstrated that these simulators 'are not just very expensive computer games', but they can in fact be used in the early learning curve, relieving a large part of the burden on patients by training in a safe and virtual environment. In **chapter 6**, the results of an international multicenter randomized controlled trial are described in which simulator-based training is directly compared to patient-based colonoscopy training. The results of a cohort study to determine the effect and usefulness of prolonged simulator-based training are described in **chapter 7**. The benefits of prolonged training are explored using patient-based performance as primary end-point as well as the point where simulator training does no longer add to improved performance.

In section four, **Patient-based endoscopy - Self-assessment and learning curves in colonoscopy**, the learning curve of novice to competent endoscopists is explored. By monitoring individual and group learning curves it becomes apparent when a trainee is not developing according to the expected learning curve, similar to growth curves for infants and children. This knowledge opens possibilities to intervene in earlier stages during training. Repetitive assessment and self-assessment of performance is helpful in determining learning objectives and exposing weak points. In **chapter 8**, the monitoring of individual and group learning curves is investigated leading to a final competency level for colonoscopy. The use of a novel developed Rotterdam Assessment Form for Colonoscopy, RAF-C, is investigated together with the potential of this tool as a method for self-improvement in a repetitive self-evaluation cycle. In **chapter 9**, the general discussion is conducted with integration of the data in conclusions and a glance to future research.

Chapter 3

Expert and construct validity of the Simbionix GI Mentor II endoscopy simulator for colonoscopy

Arjun D. Koch[1], Sonja N. Buzink[2,3], Jeroen Heemskerk[3], Sanne M.B.I. Botden[3], Roeland Veenendaal[4], Jack J. Jakimowicz[3], Erik J. Schoon[1]

[1)] *Department of Gastroenterology, Catharina Hospital Eindhoven, The Netherlands*

[2)] *Faculty of Industrial Design Engineering, Delft University of Technology, The Netherlands*

[3)] *Department of Surgery, Catharina Hospital Eindhoven, The Netherlands*

[4)] *Department of Gastroenterology, Leiden University Medical Center,The Netherlands*

Surg Endosc (2008) 22: 158–162

ABSTRACT

objectives
The main objectives of this study were to establish expert validity (a convincing realistic representation of colonoscopy according to experts) and construct validity (the ability to discriminate between different levels of expertise) of the Simbionix GI Mentor II virtual reality (VR) simulator for colonoscopy tasks, and to assess the didactic value of the simulator, as judged by experts.

methods
Four groups were selected to perform one hand–eye coordination task (EndoBubble level 1) and two virtual colonoscopy simulations on the simulator; the levels were: novices (no endoscopy experience), intermediate experienced (<200 colonoscopies performed before), experienced (200–1,000 colonoscopies performed before), and experts (>1,000 colonoscopies performed before). All participants filled out a questionnaire about previous experience in flexible endoscopy and appreciation of the realism of the colonoscopy simulations. The average time to reach the cecum was defined as one of the main test parameters as well as the number of times view of the lumen was lost.

results
Novices (N = 35) reached the cecum in an average time of 29:57 (min:sec), intermediate experienced (N = 15) in 5:45, experienced (N = 20) in 4:19 and experts (N = 35) in 4:56. Novices lost view of the lumen significantly more often compared to the other groups, and the EndoBubble task was also completed significantly faster with increasing experience (Kruskal Wallis Test, $p < 0.001$). The group of expert endoscopists rated the colonoscopy simulation as 2.95 on a fourpoint scale for overall realism. Expert opinion was that the GI Mentor II simulator should be included in the training of novice endoscopists (3.51).

conclusion
In this study we have demonstrated that the GI Mentor II simulator offers a convincing realistic representation of colonoscopy according to experts (expert validity) and that the simulator can discriminate between different levels of expertise (construct validity) in colonoscopy. According to experts the simulator should be implemented in the training programme of novice endoscopists.

INTRODUCTION

Training skills in endoscopy for diagnostic and therapeutic procedures is essential and requires a great deal of hands-on training [1]. Virtual reality (VR) simulators offer a promising option to train these skills extensively prior to training in real-life colonoscopy, without jeopardizing patients or causing them unnecessary discomfort [2]. The use of VR training prior to performing real flexible endoscopy on patients enables novice endoscopists to go through part of their proficiency curve before submitting patients to their relatively insufficient endoscopy skills. This might not only be advantageous for the patients undergoing endoscopy, but might also prevent complications and potential consequences resulting in medicolegal litigation. One of the simulators in the field of flexible endoscopy is the GI Mentor II (see Figure 1). VR simulators have been used extensively in different fields of expertise before applying these procedures to patients. In the United States of America simulator training is mandated by the Accreditation Council for Graduate Medical Education (ACGME) in laparoscopic procedures for surgical residents [3]. The first step is to validate the simulator construct properly and verify its didactic value, before implementing simulators in teaching programmes or developing a new curriculum for flexible endoscopy around them. Some studies have already been published on this subject [4–6], but the presented outcomes lacked power due to their

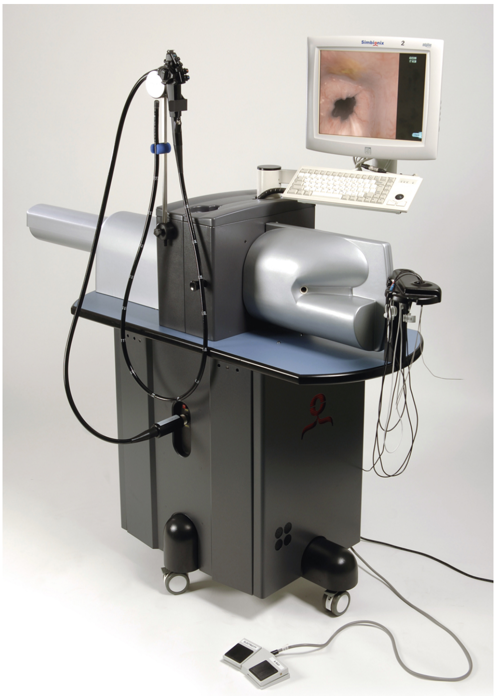

Figure 1. The GI Mentor II virtual reality simulator, the setup for training in lower endoscopy

relatively small sample sizes. In addition, some cases did not study the validity of endoscopy, but for example only the EndoBubble module, a computer simulation skills test measuring how long it takes a person to pop 20 balloons in a virtual tunnel.

OBJECTIVE

The main objectives of this study were: (1) to establish the degree of representation of real-life colonoscopy on the Simbionix GI Mentor II VR colonoscopy simulation, as judged by experts (expert validity), (2) to determine whether the GI Mentor II simulator can distinguish between various degrees of expertise in endoscopy, judged by novice, intermediate experienced, experienced and expert endoscopists performing VR colonoscopy (construct validity), and (3) to assess the didactic value of the simulator, as judged by experts.

MATERIAL AND METHODS

simulator

The simulator used in this study was the Simbionix GI Mentor II (Simbionix Ltd. Israel, software version 2.7.3.0) (Figure 1). The GI Mentor II can simulate upper GI tract endoscopies such as esophagogastroduodenoscopy, endoscopic retrograde cholangiopancreatographies, and endoscopic ultrasound. The lower GI tract endoscopies it simulates are sigmoidoscopy and colonoscopy. The simulator records a range of parameters upon each exercise, which can be used to assess performance objectively. The endoscope used is a customized Pentax ECS-3840F endoscope.

participants

Participants were allocated to four groups to assess the validity and didactic value of the GI Mentor II simulator. The first group, the novices, was defined as participants without any flexible endoscopy experience; they were all medical interns or residents. The second group was intermediate experienced, with fewer than 200 colonoscopies performed before. In the third group experienced participants all performed more than 200 colonoscopies but fewer than 1,000. The fourth group consisted of experts, all of whom had performed more than 1,000 colonoscopies. These categories were chosen based upon several other studies, the demands for Dutch accreditation for colonoscopy, and the accreditation demands of the British Society of Gastroenterology, which advocates 200 colonoscopies under supervision during training [4, 6–8]. All persons were either invited to participate within our hospital, or participated during a national congress of the Dutch Society of Gastroenterology in spring 2006. The groups consisted of at least 28 persons to ensure sufficient statistical power [9]. A

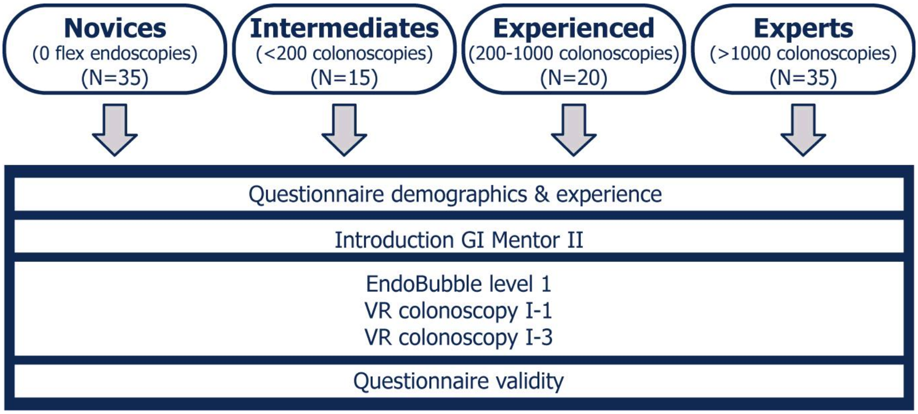

Figure 2. The study design

post hoc sample size calculation based on the results for time to finish the EndoBubble task showed a minimal sample of 26 participants in the novices group to achieve a power of 0.95. Originally, the intermediate experienced and experienced participants formed one group, but as the expertise level and performance within this group varied considerably, this groups was split. A schematic setup of the study design is presented in Figure 2.

questionnaire

All participants were asked to fill out a questionnaire on demographics and their general medical and endoscopy experience. It also included the number of endoscopies performed annually and number of years registered as a skilled professional endoscopist. After the simulator run the participants were asked to answer questions about their appreciation of the realism of the colonoscopy exercises performed. Appreciation was expressed on a four-point Likert scale [10] varying from very unrealistic (1) to very realistic (4). Questions were asked about the realism of imaging, simulator setup, endoscope control and both haptic and visual feedback. Experts were asked whether the GI Mentor II could be used as a teaching device for novice endoscopists and whether experience on the simulator could be useful in practice.

simulation modules

All participants first performed the hand–eye coordination task (EndoBubble level 1) of popping all 20 balloons in the test as quickly as possible, without touching the walls. Next, the participants performed VR case numbers 1 and 3, both from colonoscopy module 1. These cases were carefully selected for their discriminative value; both cases are straightforward colonoscopies, without any abnormalities such as polyps, tumours, or inflammation. Case number 1 is a relatively easy colonoscopy to perform, whereas case number 3 is more difficult, requiring the endoscopist to apply techniques such as straightening the endoscope during loop formation and applying torque to the endoscope shaft. The assignment given for the VR colonoscopies was to reach the cecum as quickly as possible with as little patient discomfort as possible. Patient discomfort was defined as the estimated percentage of time the virtual patient was in excessive pain and the number of times excessive local pressure was caused. Other relevant test parameters were the percentage of time spent with clear view and the number of times view of the lumen was lost. The task was considered accomplished when the cecum was reached.

data analysis

SPSS 13.0 software was used to perform descriptive statistics and Kruskal–Wallis tests for statistic analysis of the data.Aseparate analysis between groups was performed using a two-tailed Mann–Whitney exact U test. A p-value of less than 0.05 was considered significant. The data showed a nonparametric distribution, therefore the median and range of performance parameters are presented as primary values.

RESULTS

participants

Thirty-five novices, 15 intermediates, 20 experienced, and 35 expert endoscopists participated in the study. The average number of colonoscopies performed annually by experts was 445, and their mean number of years registered as a gastroenterologist was 7.7 (range 0–35 years).

construct validity

Data output by the simulator are presented in Tables 1 and 2. The EndoBubble task was completed faster by the experts and experienced endoscopists than by novices, with fewer wall collisions. These differences were statistically significant (Kruskal–Wallis test) (Table 1). Also the colonoscopy tasks were completed faster ($p < 0,001$, Kruskal–Wallis test), with less patient comfort and better visibility by experts and experienced endoscopists (Table 3). Novice endoscopists (N = 35) reached the cecum in a mean time of 29:57 (min:sec) in colonoscopy case 3, intermediate experienced (N = 15) in 5:45, experienced (N = 20) in 4:19, and experts (N = 35) in 4:56. Novices lost view of the lumen significantly more often than the other groups. A separate analysis between groups using a Mann–Whitney exact U test demonstrated no significant difference between the intermediate, experienced and expert groups on all parameters. They all completed the task faster than the novices (see Table 4).

Table 1. EndoBubble hand-eye coordination task.

Experience		Time to Finish (min:sec)	Number of times wall touched
Novice	Mean	6:56	1.9
N=35	Median	5:58	1.0
	Range	1:24 - 20:25	0 - 20
Intermediate	Mean	1:56	1.1
N=15	Median	1:41	0.0
	Range	0:54 - 4:02	0 - 5
Experienced	Mean	1:37	0.9
N=20	Median	1:21	0.0
	Range	0:43 - 5:33	0 - 9
Expert	Mean	1:24	0.3
N=35	Median	1:13	0.0
	Range	0:49 - 3:25	0 - 2
Kruskal-Wallis	Chi-Square	63.151	9.374
	Asymp. Sign.	0.000	0.025

Table 2. Colonoscopy Module 1, Cases 1 and 3.

Experience			Time to reach cecum (hour:min:sec)	% of time spent with clear view	Lost view of lumen	Excessive local pressure	% of time patient was in pain	Excessive loop formed
Case 1	Novice	Mean	6:47	96	0.4	0.5	13.3	0.83
	N=35	Median	6:16	97	0	0	11	0
		Range	1:53 - 15:08	82 - 99	0 - 3	0 - 3	0 - 44	0 - 6
	Intermediate	Mean	1:36	97	0	0	8	0.6
	N=15	Median	1:40	98	0	0	5	0
		Range	0:55 - 2:52	91 - 100	0	0	0 - 30	0 - 3
	Experienced	Mean	1:23	98	0	0.2	9.2	0.7
	N=20	Median	1:21	98	0	0	8	1
		Range	0:48 - 2:43	89 - 100	0	0 - 1	0 - 27	0 - 3
	Expert	Mean	1:23	98	0	0	14.5	1.49
	N=35	Median	1:17	98	0	0	12	1
		Range	0:42 - 3:16	94 - 100	0 - 1	0	0 - 57	0 - 10

Table 2. *Continued*

Experience			Time to reach cecum (hour:min:sec)	% of time spent with clear view	Lost view of lumen	Excessive local pressure	% of time patient was in pain	Excessive loop formed
Case 3	Novice	Mean	29:57	86	3.2	3.89	2.2	4.77
	N=35	Median	23:42	85	3	3	0	1
		Range	4:48 - 1:28:19	72 - 96	0 - 12	1 - 14	0 - 24	0 - 34
	Intermediate	Mean	5:45	89	1.1	2.1	0.9	1.13
	N=15	Median	4:21	92	1	2	0	0
		Range	2:28 - 13:41	78 - 97	0 - 4	0 - 6	0 - 4	0 - 8
	Experienced	Mean	4:19	91	0.6	1.9	1.0	1.6
	N=20	Median	3:50	91	0	1	0	1
		Range	2:27 - 7:02	73 - 99	0 - 3	0 - 8	0 - 4	0 - 9
	Expert	Mean	4:56	89	0.9	1.6	2	2.51
	N=35	Median	4:03	90	1	1	1	1
		Range	1:38 - 15:39	68 - 99	0 - 4	0 - 6	0 - 10	0 - 12

Table 3. Statistics Colonoscopy Module 1, Cases 1 and 3.

		Time to reach cecum	% of time spent with clear view	Lost view of lumen	Excessive local pressure	% of time patient was in pain	Excessive loop formed
Case 1	Chi-Square	69.043	13.889	18.415	19.783	7.101	10.691
	Asymp. Sig.	0.000	0.003	0.000	0.000	0.069	0.014
Case 3	Chi-Square	65.559	6.978	41.936	28.794	4.284	4.856
	Asymp. Sig.	0.000	0.073	0.000	0.000	0.232	0.183

Kruskal-Wallis Test

Table 4. Differences between groups Module 1, Cases 1 and 3.

		Time to reach cecum	% of time spent with clear view	Lost view of Lumen	Excessive local pressure	% of time patient was in pain	Excessive loop formed
Case 1	Novice vs Intermediate	0.000	0.177	0.039	0.013	0.070	0.743
	Intermediate vs Experienced	0.166	0.617	1.000	0.244	0.385	0.547
	Experienced vs Expert	0.962	0.621	1.000	0.043	0.077	0.020
	Intermediate vs Expert	0.141	0.259	1.000	1.000	0.018	0.009
Case 3	Novice vs Intermediate	0.000	0.104	0.000	0.004	0.584	0.040
	Intermediate vs Experienced	0.257	0.394	0.285	0.503	0.771	0.184
	Experienced vs Expert	0.969	0.297	0.153	0.942	0.154	0.726
	Intermediate vs Expert	0.326	0.757	0.870	0.416	0.111	0.090

Mann Whitney Test Exact Significance (2-tailed)

expert validity

The group of expert endoscopists rated the colonoscopy simulation 2.95 on a four-point Likert scale for overall realism. Anatomical representation was rated 2.58, and the simulator setup 3.14. Endoscope control scored 3.21. Haptic feedback was rated 2.57.

didactic value

Expert opinion was that the GI Mentor II simulator should be included in the training of novice endoscopists (3.51 on a four-point Likert scale) and that expertise gained on the simulator was considered applicable in a clinical curriculum (rated 3.29 out of 4). The simulator was not considered suitable for certification of trained endoscopists (rated 2.29 out of 4).

DISCUSSION

This study represents the largest and most detailed study on the validity of this type of colonoscopy simulator so far. The data show that the simulator can discriminate clearly between endoscopists of different expertise levels performing different colonoscopy tasks. Differences were statistically significant using relatively large sample sizes in all three exercises, the EndoBubble task as well as cases number 1 and 3. The difference between our study and previous studies by others is that we focused on the basic aspects of navigation for colonoscopy itself, rather than on the hand–eye coordination task alone, used for example in the study by Ritter et al. [4], and that we included more participants in four separate groups with different levels of expertise [4–6, 11, 12]. in this way we were able to demonstrate that the GI Mentor II can distinguish between expertise levels up to the level of an intermediate experienced endoscopist, who has performed around 200 colonoscopies. In a similar study Sedlack et al. [5] describe a limited construct for a different simulator (AccuTouch, Immersion Medical). Felsher et al. [11] demonstrated differences between novices and experts in large sample sizes but did not compare novices to intermediate levels of expertise.

In this study we have demonstrated convincing expert validity for colonoscopy on the GI Mentor II virtual simulator. This in contrast to other studies focusing on the EndoBubble task as a validation study [4] and not dealing with the subject of expert validity [4, 6, 7, 11, 12]. The colonoscopy tasks were considered as accomplished once the participants reached the cecum. Asking the participants to inspect the mucosa on the way back through the colon does not, in our opinion, provide a proper representation of the endoscopists skills in manoeuvring through the colon, as other aspects besides the basic navigation skills of the endoscopist could influence the performance parameters provided by the simulator considerably in this case. This might lead to very different end times depending, for example, on the carefulness of the endoscopist.

This study demonstrates that the GI Mentor II simulator offers a convincing, realistic representation of colonoscopy according to experts. The overall assessment was good. Expert opinion was that the simulator can be used as a teaching tool for novice endoscopists. The simulator's haptic feedback is doubtful. Inexperienced residents can be trained in the skills necessary in flexible endoscopy such as steering control, straightening the endoscope during loop formation and applying torque up to a certain level.

CONCLUSION

The current study demonstrates that the GI Mentor II simulator offers a convincing, realistic representation of colonoscopy according to experts (expert validity) and that the simulator can discriminate up to the level of intermediate experienced endoscopists (construct validity) in colonoscopy. In the cases used the simulator could not discriminate between intermediate, experienced and expert endoscopists. The next step will be a study to determine whether novice endoscopists can develop a learning curve that will actually improve their endoscopic skills applied to real patients.

REFERENCES

1. Thomas-Gibson S, Williams CB (2005) Colonoscopy training new approaches, old problems. Gastrointest Endosc Clin N Am 15(4): 813–827
2. Sedlack RE, Kolars JC, Alexander JA (2004) Computer training enhances patient comfort during endoscopy. Clin Gastroenterol Hepat 2: 348–352
3. Roberts KE, Bell RE, Duffy AJ (2006) Evolution of surgical skills training. World J Gastroenterol 12(20): 3219–3224
4. Ritter E, McClusky Dr, Lederman A, Gallagher A, Smith C (2003) Objective psychomotor skills assessment of experienced and novice flexible endoscopists with a virtual reality simulator. J Gastrointest Surg 7: 871–878
5. Sedlack RE, Kolars JC (2003) Validation of a computer-based colonoscopy simulator. Gastrointest Endosc 57(2): 214–218
6. Mahmood T, Darzi A (2003) A study to validate the colonoscopy simulator. Surg Endosc 17(10): 1583–1589
7. Datta V, Mandalia M, Mackay S, Darzi A (2002) The PreOp flexible sigmoidoscopy trainer. Validation and early evaluation of a virtual reality based system. Surg Endosc 16: 1459–1463
8. BSG Guidelines (1994) Gastro-Intestinal Endoscopy in General Practice. Gut 35: 1342
9. Field A (2005) Discovering statistics using SPSS. Sage Publications Inc.
10. Likert R (1932) A Technique for the Measurement of Attitudes. Arch Psych 140: 55
11. Felsher JJ, Olesevich M, Farres H, Rosen M, Fanning A, Dunkin BJ, Marks JM (2005) Validation of a flexible endoscopy simulator. Am J Surg 189(4): 497–500
12. Grantcharov TP, Carstensen L, Schulze S (2005) Objective assessment of gastrointestinal endoscopy skills using a virtual reality simulator. JSLS 9(2): 130–133

Chapter 4

Acquiring basic endoscopy skills by training on the GI Mentor II

S.N. Buzink[1,2], A.D. Koch[3], J. Heemskerk[2], S.M.B.I. Botden[2], R.H.M. Goossens[1], H. de Ridder[1], E.J. Schoon[3], J.J. Jakimowicz[2]

[1)]Faculty of Industrial Design Engineering, Delft University of Technology, the Netherlands

[2)]Department of Surgery, Catharina Hospital Eindhoven, the Netherlands

[3)]Department of Gastroenterology, Catharina Hospital Eindhoven, the Netherlands

Surg Endosc. 2007 Nov;21(11):1996-2003

ABSTRACT

background

Achieving proficiency in flexible endoscopy requires a great deal of practise. Virtual reality (VR) simulators could provide an effective alternative for clinical training. The objective of this study is to gain insight into the proficiency curve for basic endoscope navigation skills when training on the GI Mentor II.

methods

Thirty novice endoscopists performed four preset training sessions. In each session they performed one EndoBubble task and multiple VR colonoscopy cases (two in first session, and three in subsequent sessions). VR colonoscopy I-3 was repeatedly performed as last VR colonoscopy in each session. The assignment for the VR colonoscopies was to visualise the cecum as fast as possible, without causing patient discomfort. Five expert endoscopists also performed the training sessions. Additionally, the performance of the novices was compared with performance of 20 experienced and 40 expert endoscopists.

results

Novices progressed significantly, in particular for time to accomplish the tasks (Friedman's ANOVA $p<.05$, Wilcoxon Signed Ranks $p<.05$). Experts did not improve significantly, except for the percentage of time the patient was in excessive pain. For all runs, the performance of the novices differed significantly from those of both experienced and expert endoscopists (Mann-Whitney U, $p<.05$). Performance of the novices in the latter runs differed less from those of both experienced and expert endoscopists.

conclusions

This study demonstrates that training in both VR colonoscopy and EndoBubble tasks on the GI Mentor II improves basic endoscope navigation skills of novice endoscopists significantly.

INTRODUCTION

Navigation through the colon with a flexible endoscope is technically demanding, like many other image-based procedures [1]. It requires a high level of both psychomotor and visual-spatial skills. Consequently, trainees need a great deal of hands-on experience to master colonoscopy skills. Traditional assessment and accreditation methods are mainly based on a minimal number of supervised procedures, after which average trainees are expected to have achieved a sufficient level of proficiency. Though there are recommendations regarding these minima [2-6], the minima suggested differ considerably [7, 8]. There is a growing need for more objective methods for proficiency assessment, and a desire to training until a pre-determined level of actual proficiency instead [9-11]. In addition, training in basic endoscopy skills within a clinical setting is losing acceptance due to ethical and economic considerations [9, 11-13]. This necessitates novice endoscopists to train in the fundamentals of colonoscopy in a skills lab setting.

Virtual reality (VR) simulators could provide an effective alternative for clinical training and supply educators with objective data about the proficiency of their trainees. Currently, VR simulators are obtaining an increasingly prominent position within medical education, and they have enhanced training programmes for endoscopic skills [1, 14]. VR simulators are currently still being thoroughly evaluated, as their application must be proven valid before widespread integration in education and training programmes [12, 15, 16]. Most VR simulators record multiple performance parameters, which are assumed to provide objective insight into the proficiency level of the trainee. Some of the parameters provided by the simulators are calculated using multiple variables recorded by the simulator (see Appendix A). Currently, one of the major issues for the application of VR simulators in training programmes is to determine which types of exercises are most appropriate, and which (combination of) performance parameters represents performance best [9, 10].

Several VR simulator systems are currently available for lower gastrointestinal flexible endoscopy [12, 15, 16]. The validity of the different systems for lower gastrointestinal endoscopy skills has been studied before [7, 8, 12, 15, 17-25]. However, most of these studies did not focus on the basic tasks first, but included multiple (complex) tasks, or used relatively small numbers of participants. Based on our experience with validation of simulators for assessment and training in laparoscopic skills [16, 26], we evaluated the GI Mentor II for basic navigation skills for colonoscopy, and proved its construct validity and didactic value [20]. However, little is known about the improvement of colonoscopy skills by repetitive training on the GI Mentor II. The objective of this study is to gain insight into the first part of the proficiency curve for basic endoscope navigation skills when training on the GI Mentor II.

MATERIALS AND METHODS

This study investigates several aspects of the learning trajectory on the GI Mentor II simulator. First, the performance of 30 novice endoscopists (with no prior flexible endoscopy experience) on the simulator was assessed over a series of training sessions. Second, the performance of five expert endoscopists was investigated for the same series of training sessions. Third, the scores of novice participants were compared with those of 20 experienced (200-1000 colonoscopy procedures performed) and 40 expert (>1000 colonoscopy procedures performed) endoscopists, to assess their performance within a wider context.

participants

Thirty novices in flexible endoscopy participated in the study (10 males, 20 females, mean age 25.5 years), all of them medical interns (N=23) or residents in training (N=7). Five expert endoscopists (all male, mean age 46.2 years) also performed the four training sessions. In addition, the performance data of 20 experienced and 35 expert endoscopists, who participated in our validation study, was used [20]. The tasks performed on the simulator in the validation study were exactly the same as in the first training session of the study presented here [20].

simulator

The GI Mentor II VR simulator for flexible endoscopy (Simbionix, Ltd., Israel, software version 2.7.3.0) (Figure 1, Chapter 3) was used in this study. The GI Mentor II provides hands-on training by various modules for training in basic psychomotor endoscopy skills, lower, and upper flexible endoscopy procedures on a mannequin with a mouth and a rectal end. The endoscope used is a customised Pentax ECS-3840F endoscope. Steering and torque of the endoscope are controlled as in real endoscopy; insufflation and suction are also available. The computer simulation programme supplies visual and audio feedback, while dynamic force feedback devices inside the mannequin provide force feedback sensations, all corresponding to the selected training module and patient scenario. The patient scenarios for VR endoscopy vary in anatomy and pathology. The simulator provides objective measurements and statistics about each performance.

protocol training sessions

The proficiency of the participants in basic endoscope navigation was assessed during four preset training sessions (one per day) within five consecutive days (see Figure 2). Each participant performed one hand-eye coordination task (EndoBubble level 1) per session. Each of the training sessions also involved multiple different VR colonoscopy cases, with varying levels of difficulty, to avoid bias by training on only one patient scenario. The participants were not notified about the repetitive nature of the last VR colonoscopy in each session. VR colonoscopy I-3 was selected as repetitive exercise because of its discriminatory value; it is a

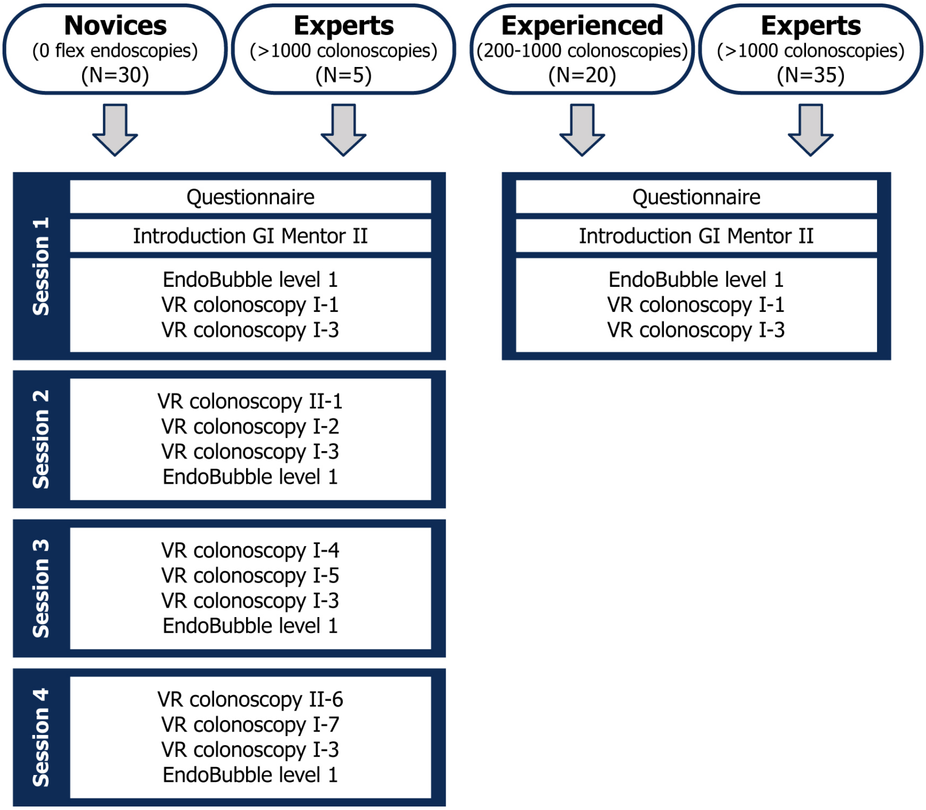

Figure 2. The study design

fairly complicated case, with a relatively winding sigmoid and a built in loop in the ascending colon and hepatic flexure [20]. The performance of the repetitive exercises within each session is defined as a run.

As the study focuses on the manipulation and navigation skills of the participants, the assignment for the VR colonoscopy exercises was to visualise the cecum as quickly as possible, with as little patient discomfort as possible. When the participant reached the cecum, the exercise was considered accomplished. Participants were instructed not to identify or treat the pathologies presented in the cases. No feedback on performance was given, other than produced by the simulator in full screen mode. Prior to the first session, the participants filled out a questionnaire on demographics, and their general medical and endoscopy experience. Next, they received an introduction about the simulator and an explanation on how to operate the controls and steer the endoscope tip. The tour and explanation were given by the researcher, following a preset objective procedure. Questions concerning the functioning of controls were answered, whenever asked during the training sessions, but no instructions were given on how to optimise performance. The participants were informed about the parameters recorded by the simulator and their scores were shown to them after each exercise. All participants performed the tasks on the simulator single-handed, without nurse-assistance for scope insertion.

performance parameters

A broad range of variables is recorded by simulator; however, time to accomplish the assignment (*time to reach the cecum* in VR colonoscopies or *time to finish the EndoBubble task*) and the *(estimated) percentage of time the patient was in excessive pain* were considered key parameters for this study (see Table 1). The *(estimated) percentage of time the patient was in excessive pain* is a composite parameter calculated by the simulator using several pain related variables. When the pain level is above the value 0.6, excessive pain is counted. If the level remains for 15 seconds, excessive pain is recorded again. See Appendix A for a more extensive explanation of the calculation of this parameter.

Table 1. Performanc e of novice participants on the EndoBubble level 1 task and VR colonoscopy I-3 per session.

training session		EndoBubble level 1		VR-colonoscopy I-3					
		time to finish (min)	number of wall collisions	**time to reach cecum (min)**	**% of time patient in excessive pain**	% of time spent with clear view	number of occasions lost view of lumen	number of occasions caused excessive local pressure	total time colon was looped (min)
1	mean	6.9	1.3	30.6	2.5	87.0	3.1	3.9	1.1
	median	5.6	0.5	23.2	0	88.0	2.5	3.5	0.1
	min-max	2.9-20.4	0-7	8.1-88.3	0-24	72-96	1-12	1-14	0-7.4
2	mean	3.7	0.6	16.0	0.3	87.2	2.8	3.3	0.1
	median	3.5	0	12.5	0	88.0	2.0	3.0	0
	min-max	2.1-7.5	0-7	5.8-50.3	0-4	74-96	0-13	1-13	0-1.0
3	mean	3.0	0.5	9.9	0.1	89.9	1.3	3.6	0
	median	2.7	0	7.5	0	90.5	1.0	2.0	0
	min-max	1.3-6.1	0-7	3.4-30.2	0-3	77-99	0-5	0-41	0-.8
4	mean	2.7	0.5	7.1	0	91.3	1.0	2.0	0
	median	2.3	0	6.4	0	91.0	1.0	2.0	0
	min-max	1.3-5.3	0-3	3.5-15.8	0	83-98	0-4	0-6	0-0

statistical analysis

SPSS 11.0.1 software was used for statistical analysis of the data. As the samples are non-parametric, the median and range of performance parameters are presented. Means are also presented in some cases to provide a complete depiction of the data. Friedman's ANOVA test and Wilcoxon Signed Ranks tests (2-tailed significance) were used to assess potential learning effects and differences in performance within groups, while differences between groups were evaluated using the Mann-Whitney U test (one-tailed significance). All differences were considered statistically significant at p≤.05.

RESULTS

performance of novices

The novices improved their performance on both VR colonoscopy I-3 and the EndoBubble task considerably (see Table 1 and Figure 3). In the fourth run, the median *time to finish the EndoBubble task* shortened by 58.5%, while the range decreased with 77%, and the median *time to reach the cecum* for VR colonoscopy I-3 shortened by 72.3%, while the range decreased

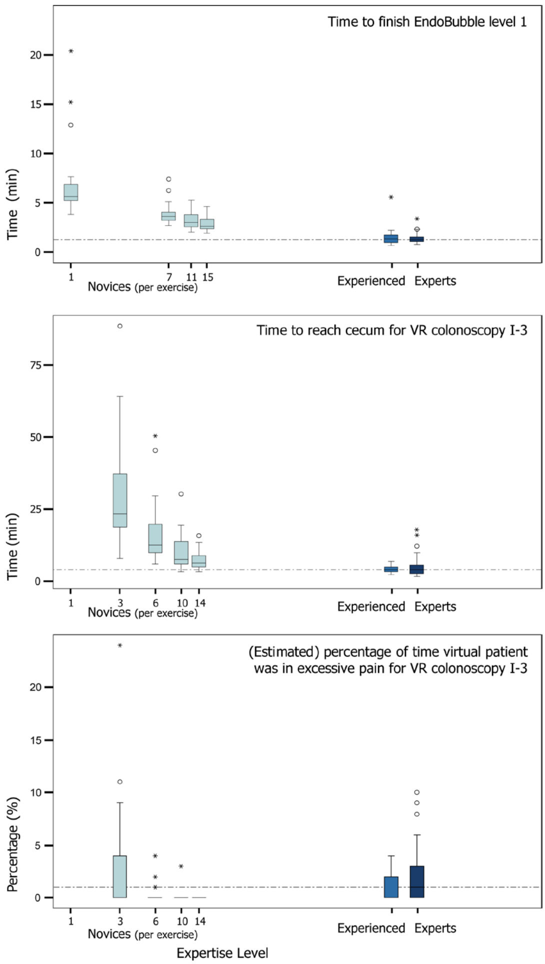

Figure 3. Boxplots for the performance of 30 novices, 20 experienced endoscopists (average of 600 colonoscopy procedures), and 40 expert endoscopists (experience with more than 1,000 colonoscopy procedures) showing time required to accomplish the EndoBubble level 1 task (a) and the VR colonoscopy I-3 (b), as well as the (estimated) percentage of time the virtual patient was in excessive pain (c). Expertise level is represented by performed exercises or colonoscopy procedures along a logarithmic scale. The reference line represents the median for experienced and expert performances. °Outliers. *Extremes in the data (cases with values of 1.5 to 3 and more than 3 box lengths from the upper edge of the box, respectively).

with 85%. The median *percentage of time the patient was in excessive pain* reached 0% in the fourth run for all participants. Friedman's ANOVA test shows that the performance of the novice participants differed significantly between the four runs (p<.05 for *number of wall collisions* in the EndoBubble task, p<.01 for all other performance parameters). The Wilcoxon Signed Ranks test shows that the *time to reach the cecum* for VR colonoscopy I-3 differed significantly (p≤.001) for all four consecutive runs (run 1 with run 2, run 2 with run 3, run 3 with run 4) (Z=-4.08, Z=-3.31, and Z=-3.60 respectively, based on positive ranks). The required *time to finish the EndoBubble task* differed significantly (p≤.001) between run 1 and run 2, and between run 2 and run 3 (Z=-4.64, and Z=-3.63 respectively, based on positive ranks). The performance of the novices in the first and fourth run of the VR colonoscopy I-3 differed significantly for all simulator parameters.

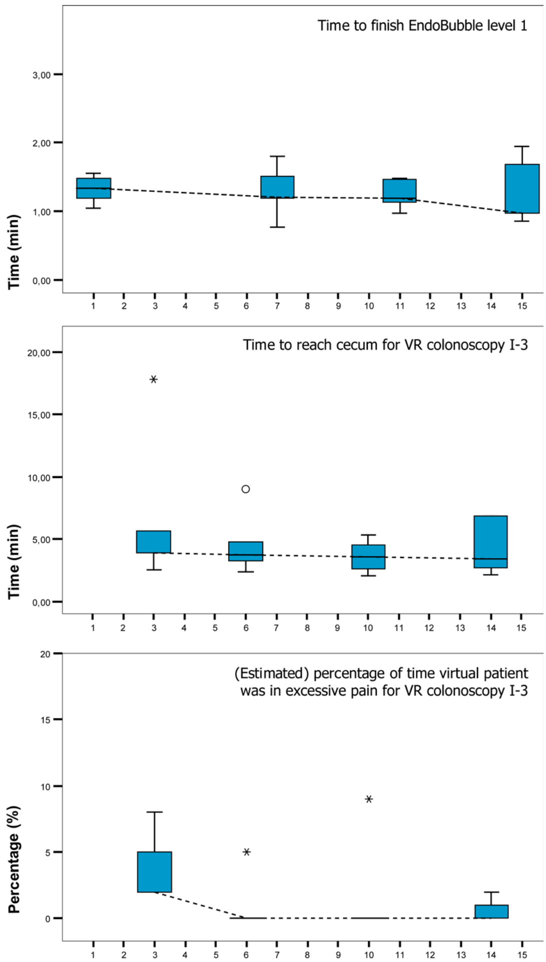

Figure 4. Boxplots of the performance parameters for repetitive exercises (EndoBubble level 1 task and VR-colonoscopy I-3) of 5 expert participants showing the time required to accomplish the EndoBubble level 1 task (a) and VR-colonoscopy I-3 (b), as well as the (estimated) percentage of time the virtual patient was in excessive pain (c).

performance of experts

At first sight, the performance of the experts appears to differ over the four runs (see Figure 4). However, Friedman's ANOVA test shows that they did not differ significantly over the four runs, except for the *percentage of time the patient was in excessive pain*. The Wilcoxon Signed Ranks test did not reveal a significant difference between their performances over the four consecutive runs either, except for the *percentage of time the patient was in excessive pain* in VR colonoscopy I-3 between the first and the fourth run.

comparing the performance of novices, experienced and expert endoscopists

The performance of novice participants in the first run and both experienced and expert endoscopists differed significantly on several parameters (see Table 2 and Figure 3). For VR colonoscopy I-3, the performance of the novices in the first run differed significantly from

Table 2. Performance of 20 experienced and 40 expert endoscopists compared with the performance of 30 novices (Mann-Whitney *U* test).

		EndoBubble level 1 task	VR-colonoscopy I-3						
		time to finish EndoBubble task (min)	number of wall collisions	**time to reach cecum (min)**	**% of time patient in excessive pain (estimated)**	% of time spent with clear view	number of occasions lost view of lumen	number of occasions caused excessive local pressure	total time colon was looped (min)
experienced endoscopists	mean	1.6	0.9	4.3	1.0	90.5	0.6	1.9	0.1
	median	1.4	0.0	3.8	0.0	91.0	0.0	1.0	0.0
	min-max	0.7-5.6	0-9	2.5-7.0	0-4	73-99	0-3	0-8	0.0-0.5
expert endoscopists	mean	1.4	0.3	4.9	2.0	88.5	1.0	1.8	0.2
	median	1.2	0.0	4.0	1.0	89.5	1.0	2.0	0.1
	min-max	0.8-3.4	0-2	1.6-17.8	0-10	68-99	0-4	0-8	0.0-1.4
novices – experienced, session 1	Z	**-5.7**	-1.6	**-5.9**	-0.3	**-2.0**	**-5.2**	**-3.8**	-1.4
	Exact Sign. (1-tailed)	**0.000**	0.056	**0.000**	0.382	**0.022**	**0.000**	**0.000**	0.079
novices – experienced, session 2	Z	**-5.3**	-0.0	**-5.8**	**-1.9**	**-1.7**	**-4.4**	**-3.0**	-1.6
	Exact Sign. (1-tailed)	**0.000**	0.521	**0.000**	**0.031**	**0.044**	**0.000**	**0.001**	0.059
novices – experienced, session 3	Z	**-4.5**	-0.5	**-4.7**	**-2.9**	-0.2	**-2.0**	-1.4	**-2.6**
	Exact Sign. (1-tailed)	**0.000**	0.293	**0.000**	**0.003**	0.424	**0.024**	0.083	**0.005**
novices – experienced, session 4	Z	**-4.2**	-0.3	**-4.0**	**-3.4**	-0.1	**-1.8**	-1.3	**-4.5**
	Exact Sign. (1-tailed)	**0.000**	0.375	**0.000**	**0.001**	0.451	**0.036**	0.101	**0.000**

Table 2. *Continued*

			EndoBubble level 1 task		VR-colonoscopy I-3					
			time to finish EndoBubble task (min)	number of wall collisions	**time to reach cecum (min)**	**% of time patient in excessive pain (estimated)**	% of time spent with clear view	number of occasions lost view of lumen	number of occasions caused excessive local pressure	total time colon was looped (min)
novices – experts	session 1	Z Exact Sign. (1-tailed)	**-7.1** **0.000**	**-2.9** **0.002**	**-7.0** **0.000**	-1.6 0.058	-1.1 0.149	**-5.3** **0.000**	**-4.6** **0.000**	-0.6 0.258
	session 2	Z Exact Sign. (1-tailed)	**-6.9** **0.000**	-0.7 0.246	**-6.3** **0.000**	**-3.4** **0.000**	-1.1 0.130	**-4.4** **0.000**	**-3.5** **0.000**	**-3.2** **0.001**
	session 3	Z Exact Sign. (1-tailed)	**-6.2** **0.000**	-0.4 0.415	**-5.2** **0.000**	**-4.3** **0.000**	-0.9 0.182	-1.3 0.106	-1.5 0.072	**-4.0** **0.000**
	session 4	Z Exact Sign. (1-tailed)	**-5.8** **0.000**	-0.9 0.187	**-3.6** **0.000**	**-4.8** **0.000**	-1.6 0.061	-0.3 0.374	-1.2 0.123	**-5.5** **0.000**

those of experienced and expert endoscopists for the *time to reach the cecum*, *occurrence of loss of view of lumen*, and *occurrence of excessive local pressure*. Additionally, the novices' performance in the first run differed from those of experienced endoscopists for *time spent with clear view*. For the EndoBubble level 1 task, the novices required a significantly longer *time to finish the task* than both experienced and expert endoscopists, and caused a higher *number of wall collisions* than expert endoscopists.

DISCUSSION

VR simulators are becoming a popular tool for training in endoscopy skills and could provide medical educators with a tool for objective proficiency assessment and an effective alternative for clinical training [9-11]. First, the use of simulators in skills-lab oriented training programmes could reduce patient discomfort and increase patient safety [9-11, 13]. Second, it could reduce workload and costs involved in experts supervising endoscopy trainees [10-12], thus improving the efficacy of the learning process. However, the construct of VR simulators and their role within training programmes are still being studied.

Our previous validation study showed that, for basic endoscope navigation, the GI Mentor II can differentiate between several levels of expertise. The GI Mentor II was also considered a valuable addition to the training programme of novice endoscopists [20]. Other studies have also established face validity and construct validity of the GI Mentor [8, 17-19]. So far, few studies have investigated learning of lower GI endoscopy skills when using the GI Mentor [7, 8, 19]. The designs of these earlier studies varied considerably in regard to the focus, tasks included, and sample size. The studies also varied considerably in training time span, type of training, and amount of training. Even though it is difficult to compare, the results encountered in our study appear to be consistent with those studies [7, 8, 19].

It is important to establish validity for all aspects of the simulator and to assess the training potential of simulators for all available training modules. It is imperative, however, to start by assessing the simulator for basic skills, in this case being the ability to navigate through the colon to the cecum. The first attribute a trainee in flexible endoscopy has to adapt to is the counter-intuitive navigation. In our study, the participants were given the assignment to reach the cecum as quickly as possible with minimal patient discomfort for the VR colonoscopies. For the EndoBubble tasks, the assignment was to pop twenty balloons as quickly as possible, while avoiding wall collisions. Assessing procedure related skills and abilities, like identification of pathologies, were intentionally not included. As the focus was on endoscope navigation, the endpoint for the VR colonoscopy exercises was reaching the cecum. For this reason, the parameters on *percentage of inspected mucosa* and *accuracy* and *efficiency of screening* were excluded.

The consistent and organised nature of the training sessions and exercises within a set time-span created a constructive environment to assess proficiency improvement: per subject, within expertise groups, and between them. It is important to minimise the influence of unfamiliarity with the simulator, or familiarity with specific cases, by using a variety of cases in each training session. After the final training session, when some of the novice participants were informed about the repetitive nature of VR colonoscopy I-3, they stated to have been unaware of this.

performance of novices

All participants improved their performance significantly over the course of four training sessions, in particular for the key parameters assessed in this study: *time to accomplish* the exercises, and *percentage of time the patient was in excessive pain*. Novices also improved their performance considerably in relation to other parameters associated with ease of navigation through the colon (like *percentage of time with clear view* and *loss of view of lumen*) and the pain level during the procedure (like *excessive local pressure* and *total loop time*) (see Table 1). This agrees with earlier studies on learning of tasks related to lower endoscopy on the GI Mentor [7, 8, 19].

The performance improvement of the trainees indicates that the difficulties often experienced by novice endoscopists when navigating through the colon with a flexible endoscope can be considerably reduced by training on the GI Mentor II. The novices appeared to be able to learn how to cope with these difficulties. As in occasions when progression of the endoscope image halted due to loop formation for example, they learned the counter-intuitive response of pulling back the endoscope shaft to progress further into the colon, most likely by trial-and-error. The considerable decrease of the pain-related parameters could entail that they also gained understanding of the factors and actions that cause pain or discomfort for the patient during flexible endoscopy. This disagrees with results of studies by Mahmood and Darzi [22], and Datta et al. [24]. The larger number of exercises, combination of different types of exercises, or influence of knowledge of results could contribute to this difference.

performance of experts

The possibility that performance improves by learning tricks that work well on the simulator, but do not necessarily improve real-life colonoscopy performance, should be taken into account when studying learning of tasks on VR systems. To verify that the GI Mentor II is not just an expensive computer game, five expert endoscopists performed the same training sessions as well. The performance of expert endoscopists shows a relatively flat profile, which demonstrates that they are on the plateau of the proficiency curve according to the GI Mentor II performance parameters (see Figure 4). In addition, the performance of the experts over the four runs is not significantly different, except for the parameter on excessive pain. This indicates that the construct of the simulator provides a valid training tool for basic endoscope navigation for colonoscopy and supports validation studies based on the GI Mentor's capability to distinguish expertise levels [7, 8, 17-20].

comparing the performance of novices, experienced and expert endoscopists

The simulator is able to distinguish between performance of novices and both experienced and expert endoscopists. However, our validation study showed that the differences between the performance of experienced and expert endoscopists on the GI Mentor II simulator are not significant [20]. In the current study, the difference between the performance of novices and experienced and expert endoscopists reduced considerably after training on the GI Mentor II, as the performance of the novices improved over the four runs. The values in Table 2 show a reduction of the difference with most of the performance parameters for both experienced and expert endoscopists, except for the *percentage of time the patient was in excessive pain,* and *total time colon was looped*. For most parameters, and in particular for the *time to accomplish* the repetitive exercises, the difference remains significant over all four runs. The *percentage of time the patient was in excessive pain* and *total time colon was looped* in Table 2 show an increase of the difference instead. The novices appear to perform increasingly better on these aspects over the four runs in comparison to the experienced and expert endoscopists (see Table 1 and Table 2). The mean ranks of the Mann-Whitney U test are lower for the novices than for the experts and experienced endoscopists, except between the novices in the first run and the experienced endoscopists. The sum of ranks was lower for novices as well, compared with the experts. Hesitancy in progression of the endoscope in combination with vigilance to cause excessive pain could play a role in the relatively low *(estimated) percentage of time the patient was in excessive pain* during the performance of VR colonoscopies I-3 by the novice participants. The *total time colon was looped* is a strong factor in the equation used to calculate the composite parameter *(estimated) percentage of time the patient was in excessive pain*, and even though the procedure time shortens significantly by run four, the absolute amount of time the patient experienced excessive pain could still be increased in comparison to the pain levels in the performance of experienced and expert endoscopists.

The performance curves of the novices appeared not to have reached a plateau within four sessions, or fifteen exercises. Even though this study does provide insight into the first part of the proficiency curve for endoscope navigation by training on the GI Mentor II, it does not provide insight into the value of the GI Mentor II for training in complete colonoscopies, which also includes inspection of mucosa or performance of therapeutic interventions. This justifies the need for further studies on the potential of the GI Mentor II for assessment of and training in flexible endoscopy, studies more longitudinal by nature, and involving more complex tasks. Transfer of skills acquired on the simulator to the performance of real-life clinical colonoscopy should be studied as well.

CONCLUSIONS

This study confirms the GI Mentor II is a valid tool for training of basic flexible endoscopy navigation skills for colonoscopy. The large sample and the strong focus on basic skills sets this study apart from earlier studies. The data provided is consistent with earlier studies on this topic, even though one-to-one comparison is difficult due to differences in study designs.

In addition, this study proves that combined training in both VR colonoscopies and the EndoBubble task on the simulator has a significant effect on the performance of novice endoscopists. The results provide additional insight into and increase the knowledge about the proficiency curve for flexible endoscope navigation when training on the GI Mentor II.

acknowledgements

The authors would like to thank all participants for their engagement in this study. We would also like to thank Cees Schot for his technical support and Sören Johnson for his linguistic advice. This study was partly funded by a grant from the Scientific Fund of the Catharina Hospital Eindhoven.

REFERENCES

1. Buzink SN, Goossens RHM, Jakimowicz JJ, Schot C, de Ridder H (2006) Image-based surgical proficiency: reflection on human factors. 16th World Congress on Ergonomics of the International Ergonomics Association. Maastricht, The Netherlands, 2006.
2. Tassios P, Ladas S, Grammenos I, Demertzis K, Raptis S (1999) Acquisition of competence in colonoscopy: The learning curve of trainees Endoscopy 31: 702-706.
3. Church J, Oakley J, Milsom J, Strong S, Hull T (2002) Colonoscopy training: The need for patience (patients) Aust N Z J Surg 72: 89-91.
4. Marshall J (1995) Technical Proficiency of Trainees Performing Colonoscopy - a Learning-Curve. Gastrointest Endosc 42: 287-291.
5. Wexner SD, Garbus JE, Singh JJ (2001) A prospective analysis of 13,580 colonoscopies. Surg Endosc 15: 251.
6. Wexner SD, Eisen GM, Simmang C (2002) Principles of privileging and credentialing for endoscopy and colonoscopy. Surg Endosc 16: 367-369.
7. Eversbusch A, Grantcharov TP (2004) Learning curves and impact of psychomotor training on performance in simulated colonoscopy: a randomized trial using a virtual reality endoscopy trainer. Surg Endosc 18: 1514.
8. Ritter E, McClusky Dr, Lederman A, Gallagher A, Smith C (2003) Objective psychomotor skills assessment of experienced and novice flexible endoscopists with a virtual reality simulator. J Gastrointest Surg 7: 871-8.
9. Satava RM (2006) Assessing surgery skills through simulation. The Clinical Teacher 3: 107-111.
10. Jakimowicz JJ, Cuschieri A (2005) Time for evidence-based minimal access surgery training - simulate or sink Surg Endosc 19: 1-3.
11. Roberts KE, Bell, R.L., Duffy, A.J. (2006) Evolution of surgical skills training. World Journal of Gastroenterology 12: 3219-3224.
12. Gerson L, Dam JV (2004) Technology review: the use of simulators for training in GI endoscopy Gastrointest Endosc 60: 992.
13. Sedlack R, Kolars J, Alexander J (2004) Computer Training Enhances Patient Comfort During Endoscopy. Clinical Gastroenterology and Hepatology 2: 348-352.
14. Sedlack RE, Kolars JC (2004) Computer Simulator Training Enhances the Competency of Gastroenterology Fellows at Colonoscopy: Results of a Pilot Study. Am J Gastroenterol 99: 33.
15. Hochberger J, Maiss J, Hahn E (2002) The use of simulators for training in GI endoscopy Endoscopy 34: 727-729.
16. Carter FJ, Schijven MP, Aggarwal R, Grantcharov T, Francis NK, Hanna GB, Jakimowicz JJ (2005) Consensus guidelines for validation of virtual reality surgical simulators. Surg Endosc 19, Number 1523 - 1532.
17. Felsher JJ, Olesevich M, Farres H, Rosen M, Fanning A, Dunkin BJ, Marks JM (2005) Validation of a flexible endoscopy simulator. Am J Surg 189.
18. Grantcharov T, Carstensen L, Schulze S (2005) Objective assessment of gastrointestinal endoscopy skills using a virtual reality simulator. JSLS, Journal of the Society of Laparoendoscopic Surgeons 9: 130-133.
19. Ferlitsch A, Glauninger P, Gupper A, Schillinger M, Haefner M, Gangl A, Schoefl R (2002) Evaluation of a Virtual Endoscopy Simulator for Training in Gastrointestinal Endoscopy. Endoscopy 34: 698-702.
20. Koch AD, Buzink SN, Heemskerk J, Botden SMBI, Veenendaal R, Jakimowicz JJ, Schoon EJ (2006) Expert and construct validity of the Simbionix GI Mentor II endoscopy simulator for colonoscopy. Congress European Association for Endoscopic Surgery (Oral presentation), Berlin, Germany, 13-16 September, 2006. Will be published in the Supplement of Surgical Endoscopy 20.

21. Sedlack RE, Kolars JC (2003) Validation of a computer-based colonoscopy simulator. Gastrointest Endosc 57.
22. Mahmood T, Darzi A (2004) The learning curve for a colonoscopy simulator in the absence of any feedback: No feedback, no learning. Surg Endosc 18: 1224.
23. Mahmood T, Darzi A (2003) A study to validate the colonoscopy simulator. Surg Endosc 17: 1583.
24. Datta V, Mandalia M, Mackay S, Darzi A (2002) The PreOp flexible sigmoidoscopy trainer. Surg Endosc 16: 1459.
25. MacDonald J, Ketchum J, Williams RG, Rogers LQ (2003) A lay person versus a trained endoscopist. Surg Endosc 17: 896.
26. Schijven MP (2005) Virtual Reality Simulation for Laparoscopic Cholecystectomy - The process of validation and implementation in the surgical curriculum outlined. Volume PhD. Leiden: University of Leiden, 2005.

APPENDIX A

The *(estimated) percentage of time the patient was in excessive pain* is calculated by the simulator using several pain related parameters:

$PL = 0.1*ELP + .04*AC + 0.7*LR^2$ (equation 1)
$PL < 0.6 : EP = EP$,
$PL \geq 0.6 : EP = EP + 1$,
after each t=15 sec $PL \geq 0.6 : EP = EP + 1$

When the pain level (PL) is above the value 0.6, excessive pain (EP) is recorded. If the excessive pain remains for another 15 seconds excessive pain is recorded again. *Excessive local pressure* (ELP) is calculated when the tip of the endoscope is pushed into the colon's wall to a depth of 1.5-2cm for more than 2 seconds. The *amount of air in the colon* (AC) is a value from 0 to 1 (0 meaning no air and 1 meaning the colon is full of air). The *loop rate* (LR) is a value between 0 and 1 (0 when the colon is totally relaxed, and 1 when the colon is extremely tensed). Percentage of time patient was in pain (P_tEP) is calculated by the time the patient was in excessive pain (EP_t) divided by the total procedural time (TP_t), see equation 2.

$P_tEP = EP_t / TP_t$ (equation 2)

Chapter 5

A second generation virtual reality simulator for colonoscopy: validation and initial experience

A.D. Koch[1], J. Haringsma[1], E.J. Schoon[2], R.A. de Man[1], E.J. Kuipers[1]

[1)]Department of Gastroenterology and Hepatology, Erasmus MC - University Medical Center Rotterdam, The Netherlands
[2)]Department of Gastroenterology and Hepatology, Catharina Hospital Eindhoven, The Netherlands

Endoscopy. 2008 Sep;40(9):735-8

ABSTRACT

background and study aims

Simulators are increasingly used in training physician's skills however data on systematic evaluation of the performance of these simulators are scarce compared to their use in aviation. The objectives of this study were to determine expert validity, construct validity, and the didactic value of the novel Olympus simulator as judged by experts.

patients and methods

Participants were novices and experts. Novices had no prior flexible endoscopy experience; experts had all performed more than 1000 colonoscopies. Participants filled out a questionnaire on their appreciation of the realism of the colonoscopy exercises performed. Exercises included a dexterity exercise and a virtual colonoscopy. Test-parameters used were points acquired in the game, time to reach the cecum, maximum insertion force and patient pain.

results

Novices (N=26) scored a median of 973 points (range -118-1393), experts (N=23) scored 1212 points (range 89-1375). This difference did not reach significance (P=0,073). Experts performed virtual colonoscopy significantly faster than novices (780 vs. 220 sec, P<0.001) but used more insertion force (11.8 vs. 11.6 N, P=0.147). Maximum pain score was higher in the expert group: 86% vs. 73%. (P=0.018). Appreciation of the realism was graded 6.5 on a 10-point scale. Experts considered the Olympus simulator beneficial for the training of novice endoscopists.

conclusions

The novel Olympus simulator discriminates excellently between the measured levels of expertise. The prototype offers a good realistic representation of colonoscopy according to experts. Although software development is ongoing, the device may already be implemented in the training programme of novice endoscopists.

INTRODUCTION

Flexible endoscopy is a key diagnostic and therapeutic modality in gastroenterology. With increasing numbers of endoscopies and interventional techniques there is an increasing demand for trained endoscopists. Traditionally such training usually involves supervised hands-on training on patients with little structured instruction and dexterity training. Proficiency in advancing a colonoscope is usually obtained by practising several hundred procedures[1]. Dutch accreditation, and the accreditation demands of the British Society of Gastroenterology both require 200 colonoscopies under supervision for certification[2]. Also, there is a need for objective assessment of skills of individual trainees[3, 4]. Virtual Reality (VR) simulators offer a promising tool to teach basic dexterity skills prior to training in real-life colonoscopy[5]. Training can be given in a learning environment, without putting patients at unnecessary discomfort or risk[6-8]. For this purpose a number of VR endoscopy simulators have been developed and validation studies have been performed[6, 7, 9-12]. In Europe there is a tendency towards implementing simulator training in the curriculum for endoscopists in training[13]. In the United States simulator training is already mandated by the Accreditation Council for Graduate Medical Education (ACGME) in laparoscopic procedures for surgical residents[14]. However, data on systematic evaluation use of VR trainers or the performance of endoscopy simulator devices are scarce compared to their use in general practice. Before implementing a novel simulator into our training programme and study proficiency curves it is vital to perform a validation study to determine an expert and construct validity. Only such a study can demonstrate validity and reliability when using a simulator as a training or assessment tool.[15]

The Olympus simulator for colonoscopy (Endo TS-1, Olympus KeyMed, Southend, UK) is a novel virtual reality simulator specifically aimed at colonoscopy skills training. This simulator is specifically designed to train endoscopic maneuvering in a virtual colonic environment, solving different loop-formations and experiencing haptic or force feedback on insertion and withdrawal. Force feedback and various scores are based on calculated wall pressure. Previous simulators have been designed to display realism in graphics, mucosal similarities and abnormalities and are equipped with tools to practice different therapeutic procedures in endoscopy. Until now they lack the degree of realism on haptic feedback the present simulator offers.

OBJECTIVES

The aim of this study was (1) to establish the degree of representation of real-life colonoscopies on the Olympus Endo TS-1 colonoscopy simulator, as judged by experts (the expert validity), (2) to determine whether the simulator can distinguish between novices and expert endoscopists performing VR colonoscopy (the construct validity), and (3) to assess the didactic value of the simulator, as judged by experts.

MATERIAL AND METHODS

virtual reality simulator

The simulator used in this study was the Olympus Endo TS-1 Virtual Reality simulator for colonoscopy. The Olympus Endo TS-1 (Figure 1) is a prototype 2nd generation virtual reality simulator. It is built mainly for colonoscopy training. The simulator records different parameters during colonoscopy such as *time to reach the cecum, shaft insertion force, shaft torque, tip section force* and *patient pain* in relation to the anatomical position in the colon. All these parameters can be displayed graphically, also in relation to previous virtual colonoscopies.

Figure 1. The Olympus Endo TS-1 simulator

The simulator is equipped with a customized Olympus CF180L endoscope with virtual angulation control, insufflation, patient rotation, variable stiffness function and virtual magnetic imaging of endoscope position.

participants

The study included two groups of participants, novices and experts. The first group, the *novices*, was defined as participants without any experience in flexible endoscopy; all were medical interns or residents. The second group consisted of *experts*, endoscopists who have performed a minimum of 1000 colonoscopies in patients.

All participants were invited to the Erasmus MC Skills Training Centre to participate and were instructed by a single trainer (ADK).

questionnaire

All participants were asked to fill out a questionnaire on demographics medical training and endoscopy experience, including the number of endoscopies performed annually and the number of years registered as a skilled professional endoscopist. In addition it included questions about experience with other medical VR simulators.

After a fixed set of simulator-based exercises participants were asked to answer questions about their appreciation of the realism of the colonoscopy exercises performed. Appreciation was expressed on a visual analogue scale (VAS) ranging form 0 to 10. This grading method was chosen because of its ability to measure a subjective parameter, which is believed to range across a continuum of values in the most objective manner. This in contrast to a 4 or 5-point Likert scale where participants are required to choose one of these fixed points. Questions were asked about the practical setup, the endoscope handling, the anatomic display, the endoscope movements and its haptic feedback.

exercises

All participants first performed a standard hand-eye coordination task named "PolypExorcise". This is a dexterity exercise with the objective to *remove ghosts* from a virtual dungeon using a simulated instrument inserted in the working channel of the endoscope with penalty points for *touching the walls*, *clumsy steering, dangerous movements and time taken to complete the task*. Next the participants performed a set of colonoscopy tasks. In this colonoscopy set endoscopists encounter an N loop in the virtual sigmoid and a near gamma loop in the virtual transverse colon. The objective was to intubate the cecum as quick as possible with the least amount of discomfort to the virtual patient. During the colonoscopy tasks the simulator simultaneously displayed the endoscopic view, the patient's position, facial discomfort expression and the endoscope position as displayed by the magnetic imager in anterior-posterior and lateral view. All participants received the same instruction on how to operate the endoscope, how to use suction, inflation and rinsing and how to use the variable stiffness.

On command of the endoscopist the patient's position was changed or abdominal compression was applied.

statistical analysis

SPSS 14.0 software was used to perform descriptive statistics and a 2-tailed Mann-Whitney U Exact test for statistical analysis of the data[16]. A p-value of less than 0.05 was considered significant. The data showed a nonparametric distribution. Therefore the median and range of performance parameters are presented as primary value.

RESULTS

Between January 2007 and June 2007 twenty-six novices and 23 experts from 24 hospitals in the Netherlands participated in the study. There were 6 experts coming from 5 different university hospitals and 17 experts working in municipal hospitals. All participants completed the exercises and the evaluation. Data on the construct validity are presented in Table 1. Experts scored more points in the dexterity exercise with less penalty points than the novices. However, this difference did not reach significance (P=0.073). Experts completed the colonoscopy task faster than novices. The median time was 220 seconds (range 119- 660,) in the expert group and 780 seconds (291-1171) for novices (P=0.001). On average, experts caused a higher maximum *patient pain* score (P=0.018). There was no difference in the maximum *shaft insertion force* between experts and novices (11.8 vs. 11.6 N, P=0.147). Also, there was no difference between the two groups in the measured *shaft torque* and *tip section force*.

Table 1. Comparison of novice and expert groups: simulator output and statistical analysis.

Experience		Dexterity exercise	Time to reach cecum	Maximum shaft insertion force	Shaft torque	Tip section force	Maximum patient pain
Novices N=26	Median	973	780	11.6	0.2	1	0.73
	Range	1511	1420	9	0.05	11	0.7
	Minimum	-118	291	11	0.2	0	0.3
	Maximum	1393	1171	20	0.25	11	1
Experts N=23	Median	1212	220	11.8	0.2	1	0.86
	Range	1286	541	8.7	0.1	4	1
	Minimum	89	119	11.3	0.15	0	0
	Maximum	1375	660	20	0.25	4	1
Mann-Whitney U	Z	-1.793	-5.469	-1.449	-0.141	-0.452	-2.376
	P	0.073	0.001	0.147	0.888	0.651	0.018

Mann-Whitney U Exact test (2-tailed)

expert validity

The group of experts appreciated the colonoscopy task with a mean score of 7.2 on a 10-point scale for difficulty. The practical set-up was rated 6.9, the endoscope handling 7.6. The display of endoscope movements was rated 7.1, the haptic feedback on steering and insertion 7.0 and 7.1 respectively.

didactic value

Expert opinion was that the Olympus Endo TS-1 should be incorporated in the training of novices in endoscopy (mean score of 8.0 on a 10-point scale). The clinical applicability of experience gained on the simulator was rated 6.5.

DISCUSSION

This is the first validation study on the prototype advanced Olympus Endo TS-1 simulator. Although the software is under further development, we demonstrate that the simulator can effectively discriminate between novices and expert endoscopists. Experts completed the colonoscopy task significantly faster than novices by exerting a significantly higher maximum pain score for the virtual patient. We believe this is caused by the fact that experts are willing to use more force to resolve a difficult loop or to pass a sharp bend. The average insertion force was not measured in this study.

According to experts the simulator offers a good representation of colonoscopy with an obvious appreciation of the haptic feedback provided by the simulator. Experts agree that this novel simulator is a promising tool for future training colonoscopy.

Similar to previous studies[6, 7], we considered the colonoscopy task accomplished once the participants had reached the cecum. This provides a better representation of the proficiency of the endoscopist when focussing on manoeuvring skills. This also corresponds better with the design of this simulator in which the emphasis is put on training the navigational skills.

The Olympus Endo TS-1 simulator distinguishes itself from other simulators because it is built solely for the purpose of teaching and training navigational skills in endoscopy rather than recognition of mucosal abnormalities or solving clinical cases. This is clearly a different approach from the virtual reality endoscopy simulators that are commercially available. We therefore label this simulator a second-generation simulator.

In our experience, using the Olympus Endo TS-1 simulator in the training of over 30 novice endoscopists, the VR simulator is already of additional value in basic dexterity training. However, future studies have to show the influence of training with this simulator on the proficiency curves of endoscopists in training. Furthermore, the simulator could also proof to be of value when training more advanced endoscopists in difficult aspects of endoscopy, such as complex loop formations, and colonoscopy without sedation.

CONCLUSIONS

In this study we have demonstrated that the Olympus Endo TS-1 offers a realistic representation of colonoscopy according to experts (expert validity). The simulator can discriminate between novices and expert endoscopists (construct validity).

The validation of this simulator is an essential step before implementing it in our teaching programme and forms the basis to conduct further studies.

REFERENCES

1. Bond JH, Frakes JT. Who should perform colonoscopy? How much training is needed? Gastrointest Endosc 1999;49:657-9.
2. BSG Guidelines: Gastro-Intestinal Endoscopy in General Practice. Gut 1994;35:1342.
3. Teague R, Soehendra N, Carr-Locke D, Segal E, Nagy G, Chao W, Sakai Y. Setting standards for colonoscopic teaching and training. J Gastroenterol Hepatol 2002;17 Suppl:S50-3.
4. Grantcharov TP, Carstensen L, Schulze S. Objective assessment of gastrointestinal endoscopy skills using a virtual reality simulator. Jsls 2005;9:130-3.
5. Thomas-Gibson S, Williams CB. Colonoscopy training--new approaches, old problems. Gastrointest Endosc Clin N Am 2005;15:813-27.
6. Koch AD, Buzink SN, Heemskerk J, Botden SM, Veenendaal R, Jakimowicz JJ, Schoon EJ. Expert and construct validity of the Simbionix GI Mentor II endoscopy simulator for colonoscopy. Surg Endosc 2007.
7. Buzink SN, Koch AD, Heemskerk J, Botden SM, Goossens RH, de Ridder H, Schoon EJ, Jakimowicz JJ. Acquiring basic endoscopy skills by training on the GI Mentor II. Surg Endosc 2007;21:1996-2003.
8. Sedlack RE, Kolars JC, Alexander JA. Computer simulation training enhances patient comfort during endoscopy. Clin Gastroenterol Hepatol 2004;2:348-52.
9. Sedlack RE, Kolars JC. Validation of a computer-based colonoscopy simulator. Gastrointest Endosc 2003;57:214-8.
10. Datta V, Mandalia M, Mackay S, Darzi A. The PreOp flexible sigmoidoscopy trainer. Validation and early evaluation of a virtual reality based system. Surg Endosc 2002;16:1459-63.
11. Felsher JJ, Olesevich M, Farres H, Rosen M, Fanning A, Dunkin BJ, Marks JM. Validation of a flexible endoscopy simulator. Am J Surg 2005;189:497-500.
12. Mahmood T, Darzi A. A study to validate the colonoscopy simulator. Surg Endosc 2003;17:1583-9.
13. Ahlberg G, Hultcrantz R, Jaramillo E, Lindblom A, Arvidsson D. Virtual reality colonoscopy simulation: a compulsory practice for the future colonoscopist? Endoscopy 2005;37:1198-204.
14. Roberts KE, Bell RL, Duffy AJ. Evolution of surgical skills training. World J Gastroenterol 2006;12: 3219-24.
15. Carter FJ, Schijven MP, Aggarwal R et al. Consensus guidelines for validation of virtual reality surgical simulators. Surg Endosc 2005;19:1523-32.
16. Field A. Discovering statistics using SPSS. Sage Publications Inc., 2005.

Chapter 6

Training and transfer of colonoscopy skills: a multinational, randomized, blinded, controlled trial of simulator versus bedside training

Adam Haycock[1], Arjun D. Koch[2], Pietro Familiari[3], Foke van Delft[4], Evelien Dekker[4], Lucio Petruzziello[3], Jelle Haringsma[2], Siwan Thomas-Gibson[1]

1) Wolfson Unit for Endoscopy, St. Mark's Hospital, Imperial College London, UK

2) Department of Gastroenterology and Hepatology, Erasmus MC - University Medical Center Rotterdam, The Netherlands

3) Digestive Endoscopy Unit, Gemelli University Hospital, Universita` Cattolica del Sacro Cuore, Rome, Italy

4) Department of Gastroenterology and Hepatology, Academic Medical Center, Amsterdam, The Netherlands

Gastrointest Endosc. 2010 Feb;71(2):298-307

ABSTRACT

background

The Olympus colonoscopy simulator provides a high-fidelity training platform designed to develop knowledge and skills in colonoscopy. It has the potential to shorten the learning curve to competency.

objective

To investigate the efficacy of the simulator in training novices in colonoscopy by comparing training outcomes from simulator training with standard patient-based training.

design

Multinational, multicenter, single-blind randomized controlled trial

setting

Four academic endoscopy centers in the United Kingdom (UK), Italy and the Netherlands.

participants and intervention

36 novice colonoscopists were randomized to 16 hours simulator training (subjects) or patient-based training (controls). Participants completed three simulator cases before and after training. Three live cases were assessed following training by blinded experts.

main outcome measurements

Automatically recorded performance metrics for the simulator cases and blinded expert assessment of live cases using Directly Observed Procedure Score (DOPS) and Global Score sheets.

results

Simulator training significantly improved performance on simulated cases compared to patient-based training. Subjects had higher completion rates ($p=0.001$) and shorter completion times ($p<0.001$) as well as demonstrating superior technical skill (reduced simulated pain scores, correct use of abdominal pressure and loop management). On live colonoscopy, there were no significant differences between the two groups.

limitations

Assessment tools for live colonoscopies may lack sensitivity to discriminate between relative novices.

conclusions

Novices trained on the colonoscopy simulator matched the performance of standard patient-based colonoscopy training and demonstrated superior technical skills on simulated cases. The simulator should be considered as a tool to develop knowledge and skills prior to clinical practice.

INTRODUCTION

Colonoscopy is a difficult procedure and standards of performance vary considerably[1-4]. Proficiency has previously been obtained by supervised hands-on training on real patients, often requiring several hundred procedures[5-7]. The ASGE currently recommends a minimum of 140 procedures before competency can be assessed[8]. This 'apprenticeship model' does not allow for standardization of training, repetition of procedures, practice, or errors. Simulated colonoscopy, if accurately modeled, would provide a training platform allowing all of these benefits without risk to patients. Despite a steady progression in computer simulation, accurately reproducing reality in colonoscopy has proven difficult[9].

The Olympus colonoscopy simulator (Endo TS-1; Olympus Keymed, Southend, UK) has been specifically designed to model aspects of colonoscopy that determine the look and feel of a real examination, including shaft looping, tip contact, insufflation or deflation of air, variable shaft stiffness, application of abdominal pressure and movement of the patient. It also provides a simulated 3-D endoscope imager view identical to that provided by ScopeGuide™. The simulator has been designed to provide a range of learning and training experiences. Two previous studies have demonstrated face, construct and expert validity[10, 11].

The primary aim of this study was the educational evaluation of the simulator as a tool for training novices in colonoscopy. The main objectives were to i) investigate performance outcomes of simulator training compared to standard patient-based training for novice colonoscopists and ii) investigate transfer of skills from simulator training to real-life colonoscopy.

MATERIALS AND METHODS

This was a prospective, multicenter, randomized, blinded evaluation of training at four endoscopy training centers in three countries; St Mark's Hospital, London, UK; Erasmus Medical Centre, Rotterdam and the Academic Medical Centre, Amsterdam, the Netherlands; A. Gemelli University Hospital, Rome, Italy. The study was approved by all four institutional review boards and was given a rating of educational evaluation by the UK National Research Ethics Service and the Amsterdam ethical committee and approval by the Rotterdam and Rome ethical committees.

participants

Participants were identified at each center by a sub-investigator. Trainees could be of any medical background (physicians, surgeons, nursing) and in a position recognized by the training institution as appropriate for training in colonoscopy. Participants were excluded if they had experience of more than 25 previous colonoscopies or flexible sigmoidoscopies, had previously attended an intensive colonoscopy training course or were previously a subject in a colonoscopy training or simulator training study. Participants who had performed more than 10 laparoscopic surgical procedures were also excluded as it was felt that they may have to 'unlearn' endoscopic skills that may conflict with gaining flexible endoscopy skills.

The participants were enrolled by a sub-investigator and randomized into subjects (simulator training) and controls (patient-based training) by the lead investigator using a computer generated block randomization protocol with 8 per block. Participants, sub-investigators and trainers in each institution were not blinded to the group allocation.

pre-training

questionnaire

A questionnaire was completed by all participants providing data on demographics, previous endoscopic experience, self-reported measures of competence and confidence in colonoscopy and expectations regarding training outcomes from the study. All participants received a standardized tutorial on the fundamentals of colonoscopy to ensure a minimum background knowledge with respect to the basic concepts of colonoscopy and scope handling.

simulator pre-training assessment

All participants performed three previously validated cases on the simulator to provide a measure of their baseline performance. Each case had a time limit of 20 minutes to intubate to the cecum, and provided a simulated ScopeGuide™ 3-D imager view to standardize it with the patient-based assessments following training. All procedures were recorded automatically by the simulator and evaluated using computer-generated parameters that had been demonstrated to have construct validity in previous studies[10, 11] (completion time, number of loops formed, time to loop resolution and insertion force) and that were considered by the investigators to have clinical importance (maximum scope insertion depth, maximum scope force applied, degree of patient discomfort).

training

simulator training
Participants allocated to the subject group received 16 hours of a standardized simulator-training program. Trainers were expected to provide minimal tuition and feedback. The training package included knowledge and skill-based learning with formative assessments in a multi-media environment and incorporated a simulated 3-D imager view. It was structured in a sequential fashion to introduce the skills and knowledge needed to progress from rectum to caecum.

patient-based training
Participants allocated to the control group received 16 hours training in patient based colonoscopy (4 half-day sessions) by an expert trainer using a ScopeGuide™ imager. During these four lists the participants were required to perform a minimum of 8 colonoscopies under one-to-one supervision. Recommendations were made for topics to be covered aiming to standardize the training. All trainees were taught to use single-handed, one person technique for colonoscopy, but the instructors were otherwise told to provide the usual training for a novice colonoscopist.

post-training assessment

simulator assessment
All participants subsequently performed the same three standardized simulator cases as for the pre-training assessment to investigate the changes in performance following training. All procedures were recorded and evaluated using the same parameters.

patient-based assessment
All participants performed three patient-based colonoscopies following their training. Cases were specifically chosen from patients scheduled to undergo a clinically indicated colonoscopy. In order to reflect the training standard, patients were excluded if they were older than 75 yrs or had pelvic or colonic surgery, and if they had a previous difficult colonoscopy. Standard operating procedures for the unit were followed during all examinations. Patients gave informed consent for participation. Sedation and monitoring were utilised as per standard practice in the unit. A ScopeGuide™ 3-D endoscope imager view was utilised for all the colonoscopies performed.

An expert assessor blinded to the group allocation of the trainee was present during all assessments. Assessors were asked not to provide any assistance, either verbal or practical, during the assessment, but could take over the procedure if they were concerned at any time about the safety of the procedure or the patient's well-being, if the trainee requested

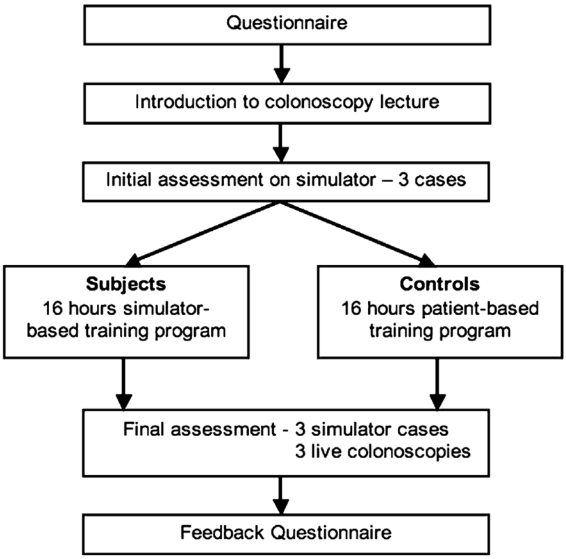

Figure 1. Flow diagram of study protocol

assistance or at any other time they considered necessary. If the assessor offered any advice or took over the procedure, then the assessment was terminated at that point. Otherwise the procedure was time limited to 20 minutes or when caecal intubation had been achieved. Caecal intubation was confirmed by the assessor by visualisation of 2 out of 3 of; the ileocaecal valve, the appendix orifice or the triradiate fold, as well as an imager view compatible with the tip of the endoscope in the caecum.

If the trainee did not achieve caecal intubation by 20 minutes, the assessor took over the procedure. The assessor straightened the scope as much as possible to remove any loops, and then recorded depth of insertion in centimetres and estimated the position of the tip using the ScopeGuide™ image. If a case was terminated due to patient factors (poor bowel preparation, extensive diverticulosis, impassable stricture or poor patient tolerance), the trainee was given the opportunity to repeat the procedure on a second suitable patient. However, no second opportunity was allowed if the termination was due to the trainee experiencing difficulty or placing the patient at risk. At the end of the case, the expert assessor made a judgment as to the difficulty of the case on a 5-point scale. Figure 1 gives a summary of the study design.

feedback

All participants completed a questionnaire following the study regarding their training and assessments.

outcome measures

Assessment of proficiency on real colonoscopy was measured using previously validated structured assessment tools; the UK Joint Advisory Group (JAG) colonoscopy Direct Observation of Procedural Skills (DOPS) assessment form[12] and an expert global rating of performance adapted for colonoscopy, the Global Score[13, 14]. Primary outcome measures were the expert assessor's score on the JAG DOPS and the Global Score. Secondary outcome measures were the time to completion, depth of insertion, and improvement in performance parameters at simulated colonoscopy.

Data from the simulator assessments were automatically exported in a coded fashion into an Access database on a secure server. Data from the patient-based assessments were entered in a coded fashion into an Excel spreadsheet.

statistical analysis

Similar studies in skills training and transfer have demonstrated effect sizes between 1 and 1.92 standard deviations when comparing treatment and control groups[14, 15]. Using the Global Score as the primary measure, it was estimated that the standard deviation of the scores would be 5 units. The study was powered to detect a 5-unit (one standard deviation) difference between the two groups. With a 5% significance level and 80% power, 16 subjects were required for each group, 32 in total. To allow for some drop out of subjects and errors in data collection, recruitment aimed to enroll a total of 36 subjects into the study.

Statistical analysis was performed on an intention-to-treat basis. For categorical data, Fisher's exact test was used to compare the results between the two groups. For ordinal data, a two-sample t-test was used to compare variables that were normally distributed, and the Mann-Whitney and Kruskal-Wallis tests for variables that were not normally distributed. For both the simulator assessments and the patient assessments, each case was analysed separately and then as a combined score from all 3 cases, which is the data presented in this paper. Logistic regression was used to investigate the change in scores from pre to post test for the simulator assessments. A p-value of less than 0.05 was considered significant.

RESULTS

study participants

Forty trainees were randomized, with thirty-six completing the study. Two trainees did not start due to limitations in availability of lists, one trainee completed the simulator pre-training assessment but had to leave for personal reasons before commencing the training, and one trainee completed the training and simulator assessments, but did not complete all three patient-based assessment cases (Figure 2).

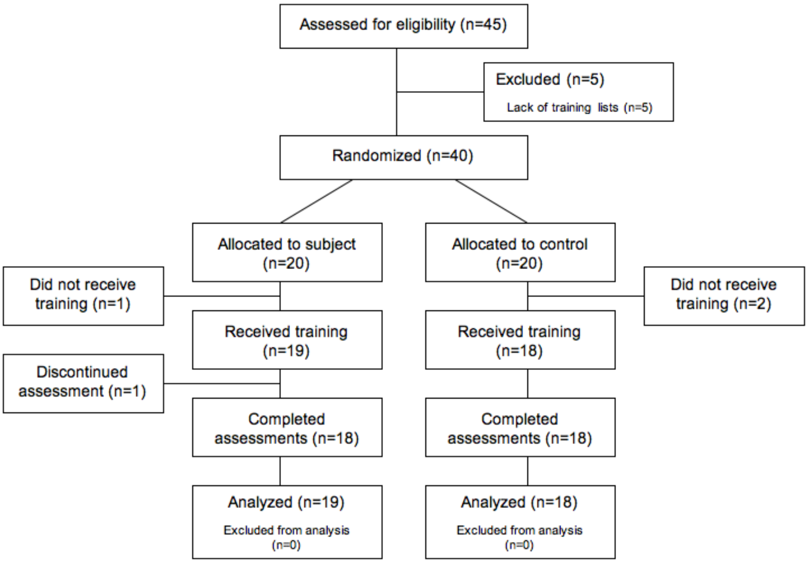

Figure 2. Flow diagram of participants through the study

Table 1. Demographics and pre-course endoscopy experience.

Variable	Controls (N=18) Median (IQR) or Number (%)	Subjects (N=19) Median (IQR) or Number (%)	P-value
Age	31 (26-33)	28 (26-30)	0.32
Male	10 (56%)	6 (32%)	0.19
Grade			
Nurse	3 (17%)	3 (16%)	0.25
General Trainee	6 (33%)	10 (52%)	
Specialist in training	8 (45%)	3 (16%)	
Other	1 (5%)	3 (16%)	
Colonoscopies witnessed	45 (21.25-137.5)	15 (7.5-125)	0.15
Colonoscopies assisted	1 (0-30)	0 (0-4)	0.21
Colonoscopies performed	0 (0-0)	0 (0-0)	0.72
Sigmoidoscopies	0 (0-1)	0 (0-0)	0.16
EGD	0 (0-2)	0 (0-0)	0.09

The subjects and controls were reasonably well matched in their demographics and previous endoscopy experience (Table 1), with no trainee having any practical colonoscopy experience. Although there were more specialists in training in the control group than the subject group, this was not statistically significant. There were some differences in the performance metrics on the simulator pre-training assessment between the groups, with the controls performing significantly better than the subjects on several measures (Table 2).

Table 2. Simulator pre-training assessment.

	Controls Median (IQR) or Number (%)	Subjects Median (IQR) or Number (%)	P-value
General			
Intubated to caecum	32 (56%)	15 (26%)	**0.002**
Maximum tip position	40 (30-40)	29 (26-40)	**<0.001**
Time taken	1032 (749-1200)	1200 (1194-1200)	**<0.001**
Technical			
Patient pain – maximum	0.63 (0.38-0.96)	0.58 (0.45-0.70)	0.77
Insertion length with embedded tip	0.10 (0.06-0.22)	0.14 (0.07-0.29)	0.51
Insertion length with obscured lens	0.04 (0.01-0.25)	0.13 (0.03-0.53)	0.05
Insertion force – maximum	13.1 (3.2)	12.2 (4.7)	0.23
Correct use of abdominal pressure	19 (36%)	13 (27%)	0.40
Correct use of variable stiffness	17 (32%)	13 (27%)	0.67
Excessive inflation	16 (28%)	18 (32%)	0.84
Looping			
Number of sigmoid loops during sigmoid intubation	1 (1-1)	1 (1-1)	0.19
Time with sigmoid loop during sigmoid intubation (s)	104 (40-182)	101 (11-194)	0.90
Number of sigmoid loops after sigmoid intubation	1 (1-1)	1 (1-2)	0.05
Time with sigmoid loop after sigmoid intubation (s)	141 (63-366)	135 (9-294)	0.20
Number of transverse loops	1 (0-1)	0 (0-1)	**0.006**
Time with transverse loop (s)	64 (0-184)	0 (0-69)	**0.01**
Time to resolve alpha loop (s)	311 (216-566)	368 (233-880)	0.34

post-training simulator assessment

Data from the simulator post-test is shown in Table 3. The subjects were significantly more likely to complete intubation to the caecum (95% vs 70%, p=0.001) and took approximately half as long to do so (407 vs 743s, p<0.001). They demonstrated superior technical skill in terms of reduced maximum patient pain scores (p=0.001), and shorter distances pushing with either an embedded tip (p<0.001) or obscured lens (p<0.001). They were more likely to use abdominal pressure correctly to assist intubation (79% vs 52%, p=0.003), but there

Table 3. Simulator post-training assessment.

	Controls Median (IQR) or Number (%)	Subjects Median (IQR) or Number (%)	P-value
General			
Intubated to caecum	38 (70%)	54 (95%)	**0.001**
Maximum tip position	40 (35-40)	40 (40-40)	**<0.001**
Time taken (s)	743 (504-1200)	407 (327-504)	**<0.001**
Technical			
Patient pain – maximum	0.45 (0.19-0.68)	0.24 (0.05-0.43)	**0.002**
Insertion length with embedded tip	0.07 (0.05-0.12)	0.03 (0.02-0.08)	**<0.001**
Insertion length with obscured lens	0.03 (0.00-0.14)	0.001 (0.00-0.005)	**<0.001**
Insertion force – maximum	13.1 (2.8)	11.8 (1.6)	**0.003**
Correct use of abdominal pressure	28 (52%)	45 (79%)	**0.003**
Correct use of variable stiffness	21 (39%)	20 (35%)	0.70
Excessive inflation	13 (24%)	7 (12%)	0.14
Looping			
Number of sigmoid loops during sigmoid intubation	1 (1-1)	1 (1-1)	0.97
Time with sigmoid loop during sigmoid intubation (s)	68 (31-104)	34 (9-61)	**0.002**
Number of sigmoid loops after sigmoid intubation	1 (1-2)	1 (1-1)	**<0.001**
Time with sigmoid loop after sigmoid intubation (s)	113 (45-240)	33 (2-68)	**<0.001**
Number of transverse loops	1 (1-1)	1 (1-1)	0.98
Time with transverse loop (s)	61 (17-168)	46 (6-81)	0.12
Time to resolve alpha loop (s)	210 (158-383)	176 (106-224)	0.11

Table 4. Logistic regression of scores from pre-training to post-training assessments.

	Odds Ratio (95% CI) (Subjects/Controls)	Mean Difference (95% CI) (Subjects/Controls)	P-value
General			
Intubated to caecum	13.3 (3.1, 53.5)		**<0.001**
Maximum tip position		3 (2, 5)	**<0.001**
Time taken		-390 (-499, -281)	**<0.001**
Technical			
Patient pain – maximum		-0.16 (-0.26, -0.06)	**0.003**
Insertion length with embedded tip		-0.07 (-0.13, -0.02)	**0.006**
Insertion length with obscured lens		-0.22 (-0.36, -0.08)	**0.002**
Insertion force – maximum		-1.2 (-2.1, -0.4)	**0.005**
Correct use of abdominal pressure	5.15 (1.93, 13.7)		**0.001**
Correct use of variable stiffness	0.94 (0.42, 2.14)		0.89
Excessive inflation	0.42 (0.14, 1.16)		0.09

Table 4. *Continued*

	Odds Ratio (95% CI) (Subjects/Controls)	Mean Difference (95% CI) (Subjects/Controls)	P-value
Looping			
Number of sigmoid loops during sigmoid intubation		-0.1 (-0.5, 0.4)	0.82
Time with sigmoid loop during sigmoid intubation (s)		-49 (-77, -22)	**0.001**
Number of sigmoid loops after sigmoid intubation		-0.7 (-1.3, 0.0)	**0.04**
Time with sigmoid loop after sigmoid intubation (s)		-98 (-146, -51)	**<0.001**
Number of transverse loops		0.5 (-0.8, 1.0)	0.09
Time with transverse loop (s)		-37 (-75, 2)	0.06
Time to resolve alpha loop (s)		-85 (-164, -7)	**0.03**

were no differences in the correct use of variable stiffness. The subjects straightened sigmoid loops more quickly than the controls (34 vs 68s, $p=0.002$) and kept the scope straighter once the sigmoid had been passed ($p<0.001$). There were no differences in the management of transverse or alpha loops. Logistic regression confirmed that the subjects improved their performance from their pre-training assessment significantly or nearing significance compared to the controls on the majority of measures (Table 4).

patient-based assessment

Two cases performed by one subject and one control were deemed to be unsuitable because of patient-related factors and were not included in the analysis. Each of the trainees performed a further case which was included. One subject was stopped by the expert assessor during their first patient-based assessment due to concerns about patient safety and did not perform any further cases. The remainder of the cases were comparable in difficulty between

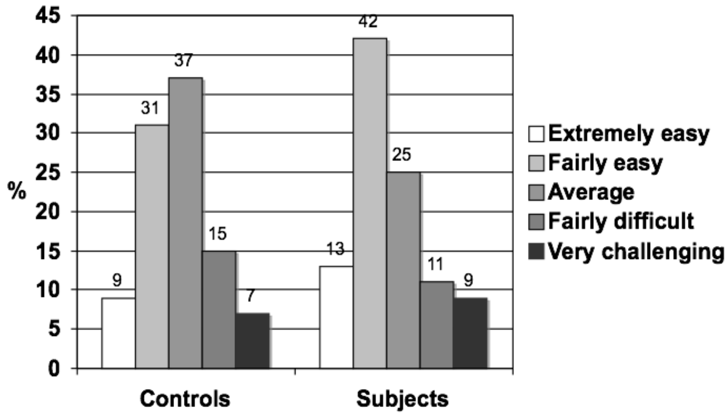

Figure 3. Distribution of case difficulty for patient-based assessments

Table 5. Results of patient-based assessments.

	Controls	Subjects	P-value
Completion of case	4 (7%)	6 (11%)	0.51
Maximum tip position			
Sigmoid	28 (52%)	29 (54%)	
Descending	12 (22%)	8 (15%)	
Transverse	8 (15%)	11 (20%)	0.73
Ascending	2 (4%)	0 (0%)	
Caecum	4 (7%)	6 (11%)	
Time taken (min)	20 (20, 20)	20 (19, 20)	0.11
Straight insertion depth (cm)	52 (21)	48 (23)	0.35
JAG DOPS	18 (14, 21)	16 (14, 22)	0.92
Global Score	17 (14, 19)	16 (14, 19)	0.35

JAG DOPS – UK Joint Advisory Group on Gastrointestinal Endoscopy Direct Observation of Procedural Skills.
Tip position expressed as number (%).
Time taken, DOPS score and Global Score as median (IQR).
Straight insertion depth as mean (SD).

Table 6. Results from the feedback questionnaire. Answers are on a 10-point Visual Analogue Scale (VAS) and expressed as Median (IQR).

	Control	Subject	P-value
Overall, how good was your training?	8.0 (6.0-9.75)	8.0 (7.0-8.5)	0.93
Overall, how useful was your training?	8.5 (8.0-10.0)	8.0 (7.0-9.0)	0.13
Overall, how enjoyable was your training?	9.0 (8.3-10.0)	8.0 (7.0-10.0)	0.18
How well do you think your training prepared you for the assessments?	7.0 (5.3-9.0)	6.0 (6.0-8.0)	0.40
How difficult did you find the simulator assessments?	6.0 (3.3-8.0)	5.0 (5.0-6.5)	0.48
How stressful did you find the simulator assessments?	5.0 (4.3-7.0)	4.0 (3.0-7.0)	0.25
How difficult did you find the patient assessments?	8.0 (7.3-8.0)	8.0 (5.5-9.0)	0.96
How stressful did you find the patient assessments?	7.0 (5.3-8.5)	6.0 (5.0-8.0)	0.28

the two groups (p=0.62), with a normal distribution for both groups (Figure 3). There were no significant differences between the groups in terms of case completion, maximum tip position achieved, time taken, straight insertion depth, JAG DOPS score or Global Score (Table 5).

feedback

The participants from both groups rated their training experience highly, with a median score of 8.0 out of 10 for both types of training (Table 6). Of note, the subjects rated training on the simulator as useful (8.0/10) and as enjoyable (8.0/10) as the controls rated training on real patients. Both groups felt reasonably well prepared for their assessments (p=0.40), although they found the patient assessments to be difficult (8.0/10).

DISCUSSION

In this study we aimed to address weaknesses in previous studies by assessing a large number of trainees from multiple centers in a prospective, randomized, blinded study. The use of patient-based training as a control allowed performance comparisons for 111 simulated and 109 clinical cases. We found that the equivalent time spent on the simulator resulted in significantly greater improvements in performance on the simulator tasks than traditional patient-based training. We also found that simulator training produced equal performance outcomes on real-life cases, demonstrating a high degree of skills transfer from the simulator to real colonoscopy. Trainees found simulator training to be useful, highly enjoyable and prepared them well for their first patient-based colonoscopies.

Most previous studies have been relatively small[16-18]. The largest, a randomized controlled multicenter trial[19], followed 45 fellows randomized to either 10 hours of simulator training prior to starting colonoscopy or to no simulator training, and found benefits from simulator training up to 80 procedures. In contrast, the only trial using patient-based training as a comparator looked at the simpler procedure of flexible sigmoidoscopy rather than colonoscopy[20], and found that the traditionally (patient-based) trained group outperformed the purely simulator trained group in most measures.

One limitation of this study is that despite randomization, there were some large but non-statistically significant differences in the level of trainee between the groups. 45% of the control group were comprised of specialists in training (defined as trainees who are currently in a formal training programme in gastroenterology), compared to 16% of the subject group. The controls had also seen more colonoscopy previously, although again this difference was not statistically significant. We speculate that although the controls did not have any increased 'hands-on' experience compared to the subjects, a specific interest in gastroenterology and increased exposure to endoscopy may have been enough to produce the differences seen in their initial simulator assessments. The statistical methods used in this study also assume that the results from the cases performed by each individual are independent of each other, which may not necessarily be the case. However, the ordinal nature and skewness of some of the data makes correcting for this possible lack of independence difficult, and would be unlikely to have a significant effect on the results.

We found that the simulator group demonstrated higher performance on the simulated cases but only equivalence on patients. We consider this may be because the assessment tools used (DOPS and Global Score) may not have been sensitive enough to discriminate between relative novices, as they are designed to assess whether a trainee is competent or not. Despite the advanced modelling achievable by the simulator, real colonoscopy is inherently more difficult and unpredictable, so novices would not be expected to achieve competence scores after only 16 hours of training. It may also be that during their training the simulator group had learned to 'play the game' and could complete the simulator cases using techniques that work on the simulator but not on patients. However, if this were so, then we would expect the subjects to perform less well on real-life colonoscopy, which was not the case. This issue is an inherent problem of even high-fidelity simulators and one of the reasons why simulation is unlikely to ever completely replace bedside training for complex procedures.

We did not include a rating of the patients' experience of the procedures, although previous work has shown that training on simulators can enhance patient comfort during colonoscopy[21]. However, one of the disadvantages of computer simulators is the difficulty in training non-technical skills such as communication and patient interaction given the lack of a 'real' environment. It is possible to combine a virtual reality simulator with a high-fidelity simulated environment[22] and with simulated patients[23], but this requires a significant expenditure of time and resources unavailable in routine clinical training. Also, unlike for sigmoidoscopy, patient ratings of colonoscopy are limited by the use of sedation and currently lack reliability and validity[24].

This study focused on novice training as it is easier to measure changes in performance on the steep slope of the learning curve. However, there are many potential advantages of simulator training for more experienced trainees. The ability to repetitively perform parts of a procedure, known as deliberate practice, has been recognised as essential to the acquisition of complex skills within medicine[25, 26]. The use of high-fidelity simulation may allow this level of practice for colonoscopy, although clinical outcome measures may need to be further validated in order to prove a significant training effect. This study also focused purely on the insertion phase, which forms only part of the skill set needed for colonoscopy. Other skills such as good withdrawal technique to achieve complete visualisation of mucosa, lesion recognition and therapeutic ability are also necessary for competency and the use of simulation to meet these training needs will need to be addressed in future studies.

The findings of this study provide strong evidence for the use of the colonoscopy simulator as a training tool for novice colonoscopists. Simulator training and standard patient-based colonoscopy training appear to be equivalent as training approaches, and use of the simulator may therefore shorten the learning curve to competency. It may be considered both as a tool to develop knowledge and skills prior to clinical practice as well as an adjunct to more traditional training methods. The ability to simulate difficult cases more realistically may be helpful in training clinicians who already possess more advanced endoscopic skills. It may also be a useful tool to assess practical skills, although further work is required to investigate its predictive value for training and credentialling.

REFERENCES

1. Harris JK, Vader JP, Wietlisbach V, Burnand B, Gonvers JJ, Froehlich F. Variations in colonoscopy practice in Europe: a multicentre descriptive study (EPAGE). Scand J Gastroenterol 2007;42(1):126-34.
2. Bowles CJ, Leicester R, Romaya C, Swarbrick E, Williams CB, Epstein O. A prospective study of colonoscopy practice in the UK today: are we adequately prepared for national colorectal cancer screening tomorrow? Gut 2004;53(2):277-83.
3. Shah HA, Paszat LF, Saskin R, Stukel TA, Rabeneck L. Factors Associated With Incomplete Colonoscopy: A Population-Based Study. Gastroenterology 2007;132(7):2297-303.
4. Radaelli F, Minoli G. Factors associated with incomplete colonoscopy: further information from a large prospective Italian survey. Gastroenterology 2007;133(4):1390-1; author reply 2.
5. Wexner SD, Garbus JE, Singh JJ. A prospective analysis of 13,580 colonoscopies. Reevaluation of credentialing guidelines. Surg Endosc 2001;15(3):251-61.
6. Bond JH, Frakes JT. Who should perform colonoscopy? How much training is needed? Gastrointest Endosc 1999;49(5):657-9.
7. Cass OW. Objective evaluation of competence: technical skills in gastrointestinal endoscopy. Endoscopy 1995;27(1):86-9.
8. Wexner SD, Litwin D, Cohen J, et al. Principles of privileging and credentialing for endoscopy and colonoscopy. Gastrointest Endosc 2002;55(2):145-8.
9. Williams CB, Thomas-Gibson S. Rational colonoscopy, realistic simulation, and accelerated teaching. Gastrointest Endosc Clin N Am 2006;16(3):457-70.
10. Haycock AV, Williams CB, Saunders BP, Thomas-Gibson S. Achieving realistic simulation: Initial validation of the Olympus Colonoscopy Simulator (ENDO TS-1). Gut 2007;56 (Suppl 3) (Suppl 3).
11. Koch AD, Haringsma J, Schoon EJ, de Man RA, Kuipers EJ. A second-generation virtual reality simulator for colonoscopy: validation and initial experience. Endoscopy 2008;40(9):735-8.
12. Colonoscopy - DOPS assessment form. The Joint Advisory Group on GI Endoscopy, 2008. (Accessed at http://www.thejag.org.uk/Forms/tabid/72/Default.aspx.)
13. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. Testing technical skill via an innovative "bench station" examination. Am J Surg 1997;173(3):226-30.
14. Park J, MacRae H, Musselman LJ, et al. Randomized controlled trial of virtual reality simulator training: transfer to live patients. Am J Surg 2007;194(2):205-11.
15. Naik VN, Matsumoto ED, Houston PL, et al. Fiberoptic orotracheal intubation on anesthetized patients: do manipulation skills learned on a simple model transfer into the operating room? Anesthesiology 2001;95(2):343-8.
16. Sedlack RE, Kolars JC. Computer simulator training enhances the competency of gastroenterology fellows at colonoscopy: results of a pilot study. Am J Gastroenterol 2004;99(1):33-7.
17. Ahlberg G, Hultcrantz R, Jaramillo E, Lindblom A, Arvidsson D. Virtual reality colonoscopy simulation: a compulsory practice for the future colonoscopist? Endoscopy 2005;37(12):1198-204.
18. Tuggy ML. Virtual reality flexible sigmoidoscopy simulator training: impact on resident performance. J Am Board Fam Pract 1998;11(6):426-33.
19. Cohen J, Cohen SA, Vora KC, et al. Multicenter, randomized, controlled trial of virtual-reality simulator training in acquisition of competency in colonoscopy. Gastrointest Endosc 2006;64(3):361-8.
20. Gerson LB, Van Dam J. A prospective randomized trial comparing a virtual reality simulator to bedside teaching for training in sigmoidoscopy. Endoscopy 2003;35(7):569-75.
21. Sedlack RE, Kolars JC, Alexander JA. Computer simulation training enhances patient comfort during endoscopy. Clin Gastroenterol Hepatol 2004;2(4):348-52.
22. Aggarwal R, Undre S, Moorthy K, Vincent C, Darzi A. The simulated operating theatre: comprehensive training for surgical teams. Qual Saf Health Care 2004;13 Suppl 1:i27-32.

23. Kneebone R, Nestel D, Wetzel C, et al. The human face of simulation: patient-focused simulation training. Acad Med 2006;81(10):919-24.
24. Redelmeier DA, Katz J, Kahneman D. Memories of colonoscopy: a randomized trial. Pain 2003; 104(1-2):187-94.
25. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med 2004;79(10 Suppl):S70-81.
26. Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. Acad Emerg Med 2008;15(11):988-94.

Chapter 7

Prolonged colonoscopy simulator training leads to improved performance during patient-based assessment

Arjun D. Koch[1], Vivian E. Ekkelenkamp[1], Jelle Haringsma[1,*)], Erik J. Schoon[2], Robert A. de Man[1], Ernst J. Kuipers[1]

[1)]Department of Gastroenterology and Hepatology, Erasmus MC - University Medical Center Rotterdam, the Netherlands

[2)]Department of Gastroenterology and Hepatology, Catharina Hospital Eindhoven, the Netherlands

[*)]current affiliation: Department of Gastroenterology and Hepatology, Maasstad Hospital Rotterdam, the Netherlands

Submitted

ABSTRACT

introduction

Virtual reality endoscopy simulators are increasingly used in the training of novice endoscopists. There are however insufficient data regarding the effect of simulator training on the early learning curve of novice endoscopists.

The aim of this study therefore was to assess the clinical performance of novice endoscopists during colonoscopy after intensive and prolonged training on a virtual reality endoscopy simulator.

methods

Eighteen trainees without any endoscopic experience were included in the study. They were divided into two groups. The simulator-training program consisted of either 50 (group I) or 100 (group II) virtual-reality colonoscopies. After 10, 30 and 50 (group II) (group I), and after 20, 60 and 100 (group II) virtual colonoscopies, trainees underwent both simulator-based (SBA) and patient-based (PBA) assessment.

results

Eighteen novices participated in the study. All completed virtual training and assessments. The mean cecal intubation time on the SBA decreased from baseline 9.50 min. to 2.20 min. at completion of the training ($P=0.002$). Colonic insertion depth during PBA improved from 29.4 to 63.7 cm ($P<0.001$). The learning effect of simulator training ceased after 60 colonoscopies.

conclusions

Virtual reality training by means of a colonoscopy simulator leads to a significant improvement of performance on the simulator itself and, more importantly, to significantly improved performances during patient-based colonoscopy. This study demonstrates the rationale for intensive simulator training in the early learning curve of novices performing colonoscopy.

INTRODUCTION

Training in procedural skills in gastrointestinal endoscopy is gradually changing from the use of threshold numbers towards a more competency based training approach. Virtual reality (VR) endoscopy simulators may be of benefit in the education of gastroenterology trainees, especially in the early learning curve. However, the major part of training remains patient-based. Virtual reality (VR) endoscopy simulators are increasingly used in the training of novice endoscopists [1-3][4-6]. This is among others due to the continued further upgrading of simulators to high levels of virtual reality [4,7-9], the introduction of competency models in medical training [10], and demands from health authorities and the public regarding physician training in general [11,12]. Training the basic endoscope navigational skills in patients only is losing acceptance.

The old adagium "See one, do one" seems no longer appropriate for educating health professionals to perform complex procedures. A recent paper defined a pre-patient training for technical skills for complex medical procedures in a virtual reality surrounding [13]. This is complementary to the ancient master-apprentice model with graduated independence and 'on the job' training. A recent Cochrane review showed that endoscopy simulators accelerate the early learning curve of novice endoscopists. There is however no convincing evidence to support superiority for either simulator-based training or patient-based endoscopy training [14].

A variety of simulators have been developed for virtual reality simulation of endoscopic procedures and interventions. These simulators provide the opportunity to familiarize endoscopists with new procedures and to repeatedly train complex procedures in order to reach a higher level of experience before performing the same procedures in patients. Numerous simulator validation studies have been performed [6,8,9,15-20], as well as studies to demonstrate increased performance of novice endoscopists after simulator training [6,12,21,22]. Remaining questions are on the optimal extension of VR training, and whether there is a point at which prolonged simulator training does not further improve performance? What is the optimal use of VR endoscopy simulator training prior to transfer to a clinical setting in a master-apprentice situation?

The aim of this study therefore was to assess the performance of novice endoscopists during patient-based colonoscopy after intensive and prolonged training on a virtual reality endoscopy simulator and to identify the point where continuation of training on the simulator ceases to have additional value on skills acquisition.

PATIENTS AND METHODS

We performed a prospective, single-centre evaluation for training colonoscopy at the Erasmus MC - University Medical Center in Rotterdam, the Netherlands. The study was approved by the institutional review board.

All participants were young physicians at the start of their training in gastroenterology. Trainees were excluded if they had any form of previous simulator or patient-based endoscopic experience. All trainees provided informed consent and completed a questionnaire providing data on demographics, previous endoscopic experience, and simulator experience.

simulator

The simulator used in this study was the Simbionix GI Mentor II (Simbionix Ltd. Israel, software version 2.7.4). The endoscope used is a customized Pentax ECS-3840F endoscope. The simulator is equipped with a training program for 20 virtual colonoscopies, two modules with each ten colonoscopies. All colonoscopy cases were randomly used for training in order to avoid bias by training only one patient scenario. Case number three of the first module was repeated after each training session for assessment of performance. This case was chosen because of its discriminative value in measuring performance as demonstrated in earlier validation studies [8,21]. It is a straightforward colonoscopy without any abnormalities such as polyps, tumours, or inflammation; however, the relatively winding sigmoid and a build in loop in the descending and transvers colon make it a fairly complicated case.

training program

All participants received a standardized tutorial on the fundamentals of colonoscopy to ensure minimum background knowledge with respect to the basic concepts of colonoscopy and colonoscope handling. An instructor (ADK or VEE) was present during the entire simulator training program. Two cohorts of participants were formed. The training program contained 20 or 10 sessions of five consecutive colonoscopies; each session ended with case three from the first module for colonoscopy. To avoid bias in the performance scores, the participants were not notified about the repetitive nature of the last VR-colonoscopy in each session. Two sessions were performed each week. This five or ten week schedule was based on the idea that distributed learning is more effective than massed training [23]. After 10, 30 and 50 (group I), or 20, 60 and 100 (group II) virtual colonoscopies participants performed two patient-based colonoscopies (Figure 1). All participants were randomly allocated to the first or second group. This elaborate division in two cohorts was chosen to minimise the learning effect of each patient-based colonoscopy performed by the trainees. As our primary aim was to assess the effect of simulator-based training on patient-based performance, we tried to limit the number of bedside colonoscopies as much as possible. The division into two cohorts led to a reduction of 50% on the total number of patient-based colonoscopies, form a

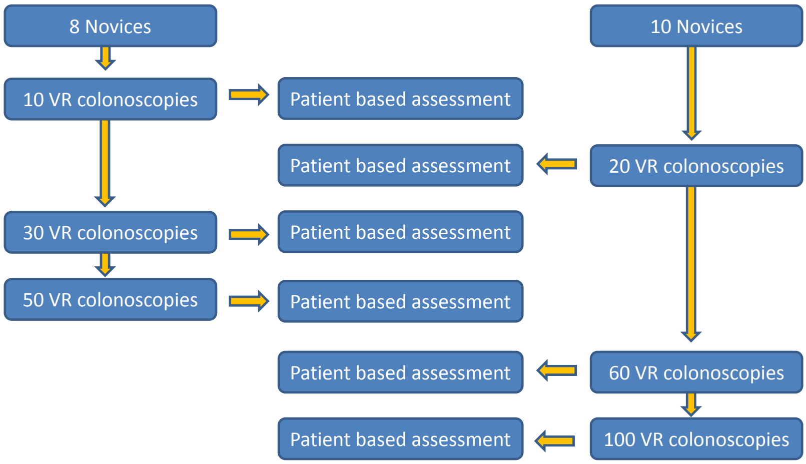

Figure 1. Training program and patient-based assessment

total of 12 to 6 procedures per participant. The simulator-based learning curve was analysed for the entire group. This means that 18 participants, both group I and II, performed 50 VR colonoscopies, while 10 of them performed an additional 50 VR colonoscopies.

patient-based colonoscopies

Patient-based colonoscopies were performed after 10, 20, 30, 50, 60 and 100 simulator colonoscopies. To minimize the impact of learning acquired by performing the patient-based colonoscopies, two actions were taken: (1) the number of patient-based colonoscopies was limited to two colonoscopies only, and (2) participants were allocated to two groups, both performing three of the total of six patient-based assessments (Figure 1). The number of two colonoscopies was consciously chosen to correct for any difficult colonoscopy for example with an inadequately cleaned colon, fixed sigmoid because of diverticulosis or previous abdominal surgery. Although ideal for statistical purposes, a greater number of real-life colonoscopies was considered to have too much impact on the simulator-derived learning curve of novice endoscopists. The mean results in terms of the number of cecal intubations and the maximum insertion depth of these two clinical colonoscopies were used for analysis.

For the same reasons, to minimize the impact on the simulator derived learning curve, the trainees were allocated to two groups with the same training program but patient-based assessments at different intervals. Colonoscopies were performed using an Olympus CF-Q160DL colonoscope. A ScopeGuide™ 3-D magnetic endoscope imager view was utilised for

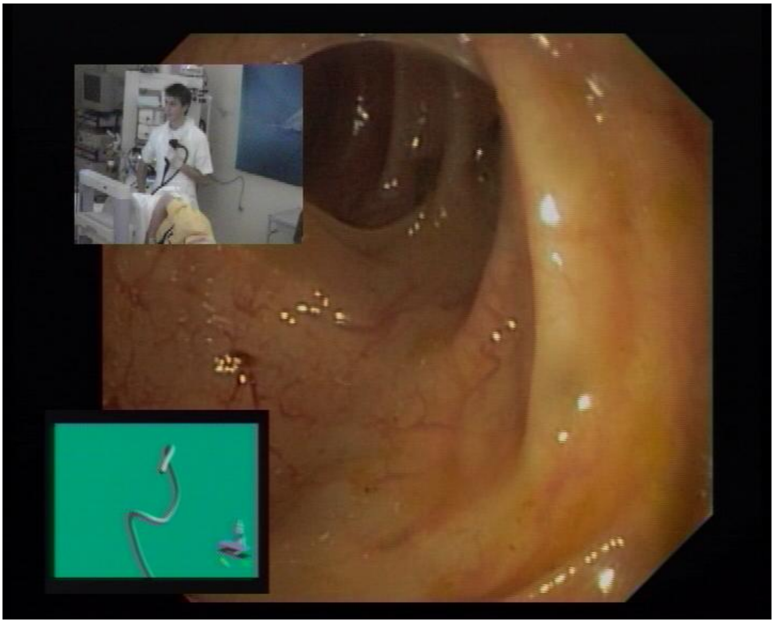

Figure 2. Tri-split video recording

all PBA procedures. Participants had access to the information on the endoscope position and loop formation, provided by the ScopeGuide™ imager. The patient-based assessments were carried out in randomly selected patients who were already scheduled for routine colonoscopy. Exclusion criteria were (i) previous colonic resections, and (ii) documented previous colonoscopies where the cecum was not reached by an expert endoscopist. An expert endoscopist was present at all times during the patient-based assessment. The expert was unaware of the number of VR colonoscopies performed by the trainee. The expert was instructed to take over the procedure (i) if he was concerned at any time about the safety of the procedure or the patient's well-being, (ii) after a fixed 20 minutes time limit, or (iii) when the trainee had reached the cecum. Patients provided informed consent for inclusion in the study. In case the trainee did not reach the cecum, the maximum insertion depth from the anal verge was measured after straightening the endoscope using the endoscopic and the ScopeGuide™ view. In case a loop was present, the endoscope was straightened while maintaining the tip of the endoscope at the point of maximum insertion as displayed by the ScopeGuide™ imager.

statistical analysis

Analyses were carried out in SAS® version 9.2. Two separate groups of novice colonoscopists followed a training curriculum of either 50 or 100 simulator-based colonoscopies. Patient-based assessment was carried out after 10, 30 and 50 simulator colonoscopies for the first group and after 20, 60 and 100 colonoscopies for the second group. This method was chosen to minimize the learning effect of each patient-based assessment. Because of the limited experience of the participants, every extra procedure was considered to possibly impact on their competence level. The two cohorts, as described, were combined and considered as one integral group for the analysis. This means that the learning curves for both groups were fused and considered as a single, group learning curve. Linear mixed models were used for analysis of the patient-based assessments of the integrated group. These models included a random intercept for trainees, in order to take the repeated measurements of the individual trainees into account. A Wilcoxon signed rank test was used for analysis of group simulator performance. Two-sided P-values less than 0.05 were considered significant. Differences between patient-based assessments at the predetermined intervals were calculated using two-tailed, paired and unpaired t-tests. Graphs and trend lines were created using standard software.

RESULTS

Eighteen trainees were included, five male and thirteen female. Eight participants were allocated to the first group, ten to the second. Their mean age was 27 years. All 18 trainees were trainees at the start of their training in gastroenterology. None of them had any previous endoscopic experience. All completed the training program and patient-based assessments (PBAs).

simulator performance

The mean cecal intubation time on the simulator for the entire group improved from baseline 9.50 min. to 2.20 min. at completion of the training (P=0.002). The learning curve is displayed in Figure 3. Eighteen participants performed 50 VR colonoscopies and only ten participants, allocated to the second group proceeded to perform 50 more VR colonoscopies to a total of one hundred. The learning curve is displayed as a mean group learning curve for the entire group. The results show a rapid improvement of performance on the simulator during the first 50 VR colonoscopies. From 60 VR colonoscopies the learning effect of prolonged training on the simulator seems to diminish.

patient-based assessments

Tables 1 and 2 show the data on the patient-based assessments for each group. The insertion depth is reported with the range and the 95% confidence interval. Also the number of cecal intubations that occurred at each session is reported. Figure 4 shows the learning curve for insertion depth during patient-based assessment related to the number of VR colonoscopies. This performance curve is displayed as a mean for the entire group.

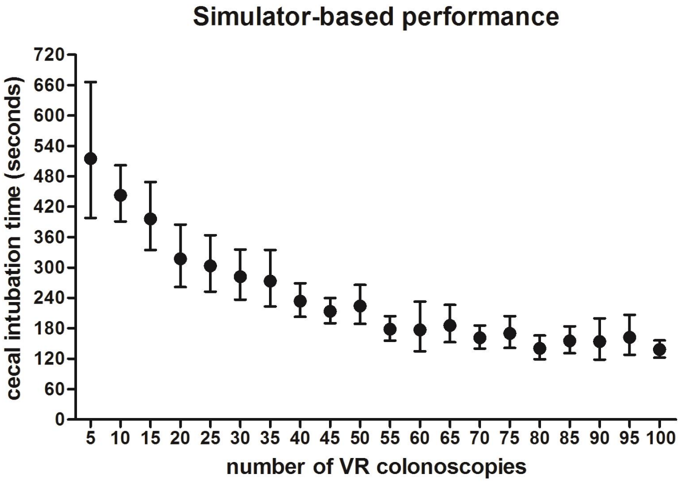

Figure 3. Simulator-based performance curve

Table 1. Performance during patient-based assessment (PBA). Group I (n=8).

Number of VR-colonoscopies	Number of PBA colon-oscopies	Insertion depth (cm)	Range (cm)	95% CI (cm)	Number of cecal intubations
10	16	29.4	12-48	22.6-36.2	0
30	16	38.6	18-70	31.4-46.1	1
50	16	58.5	23-85	50.9-66.1	1

Table 2. Performance during patient-based assessment (PBA). Group II (n=10).

Number of VR-colonoscopies	Number of PBA colon-oscopies	Insertion depth (cm)	Range (cm)	95% CI (cm)	Number of cecal intubations
20	20	36.5	18-60	30.3-42.7	0
60	20	60.5	32-95	51.8-69.1	1
100	20	63.7	25-110	55.1-72.2	3

Table 3. Differences in insertion depths.

PBA session	Mean difference in insertion	p-value*
10 vs. 100	34.3	<0.001
20 vs. 100	27.2	<0.001
30 vs. 100	25.1	<0.001
50 vs. 100	5.2	0.361
60 vs. 100	3.2	0.561

*Differences were tested with two-tailed paired t-tests within the same group and unpaired t-tests between groups.

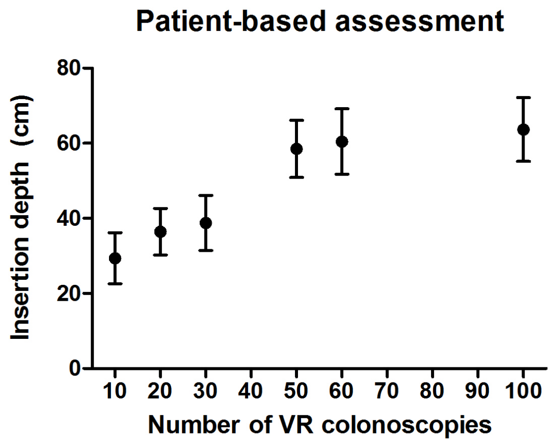

Figure 4. Patient-based assessment

Analysis of the patient-based assessments was carried out using a linear mixed model, with insertion depth as dependent variable, simulator session as predictor and a random intercept for trainees. Insertion depth increased from 29.4 cm at the first PBA to 63.7 cm during the last PBA. This increment in insertion depth over the amount of training was significant ($p<0.001$). There was no significant difference in performance between trainees ($p=0.29$). Differences between separate PBAs were calculated. The results are shown in Table 3. After 50 VR colonoscopies, a statistically significant increment in performance could no longer be demonstrated ($p=0.361$). A visual estimate of the patient-based performance curve demonstrates a plateau phase with no further improvement after approximately 60 VR colonoscopies, similar to the simulator-based performance curve.

DISCUSSION

Numerous studies have been performed to demonstrate improved performance of novice endoscopists after simulator-based training. Most of these studies show increased performance either on the simulator itself and some measure improved performance and transfer to patient-based endoscopy [24-28]. A limited number of studies have been carried out to determine the extent of simulator training versus no training. A large multicenter trial demonstrated improved competency during the first 80 patient-based colonoscopies after simulator training [29]. Another multicenter trial compared simulator-based training with traditional bedside training and demonstrated equal acquisition of skills and performance in patient-based colonoscopy [30]. A recent Cochrane review confirmed the additional value of simulator-based endoscopy training in the early learning curve of novice trainees. No superiority could be demonstrated for either simulator-based or traditional patient-based training [14].

In this study we have chosen a different approach and investigated how simulator-based training affects clinical performance and how much training is useful to still generate an increment in performance. We have demonstrated the rationale for prolonged and intensive training on a colonoscopy simulator. Continued training shows improvement of performance of novice endoscopists on the simulator itself, which is accompanied by significant improvement of performances during actual patient based colonoscopy. Prolonged simulator training provides novices with a head start for clinical endoscopy. This is thought to contribute to improvement of patient safety and diminish patient risk.

We have used both a statistical and a visual estimate to identify the point where there is no more improvement with continued training on the VR simulator. In our study, there was no longer a statistically significant increase in performance measured after 50 VR colonoscopies. However, at this point there is still a slight rise in the learning curve, indicating that there is a potential benefit for additional VR training. Visually, this point is reached after approximately 60 VR colonoscopies. From this point on there is no more improvement in simulator-based performance as well as patient-based performance. This practical approach simplifies preclinical training: it seems that when the acquisition of skills on the simulator reaches a plateau phase, the same is true for the transfer of these skills to patient-based colonoscopy. This can be translated into a training curriculum where trainees start to train their colonoscopy skills using VR simulators. When the learning curve on the simulator itself seems to reach the plateau phase, there is no need to continue VR training and the trainee can progress to "on the job" training in patient-based colonoscopy. This method replaces a threshold number of procedures by a more competence-based approach.

Another advantage of simulator training program is that it potentially diminishes instructor time. This was not measured in this study but previously reported [31]. A possible pitfall in leaving novice endoscopists completely alone during their simulator training is that 'bad habits' in handling the endoscope are developed which have to be 'unlearned' during patient-based colonoscopy. This problem was avoided in this study by providing all participants a standardized introduction on how to handle the instrument. Also during the training an instructor (ADK or VEE) was present at intervals to observe the trainees and give instructions if necessary.

Ideally, for statistical purposes, a larger number of colonoscopies had been performed per session. More patient-based assessments could have resulted in a more fitted learning curve. However, having the trainees perform more patient-based colonoscopies during ongoing simulator training would inevitably have affected the learning curve and would have biased the effect of the simulator training. In order to avoid this bias, the trainees were divided into two separate groups. Each group performed their patient-based assessments at different points in time in their training program.

Based on the findings in this study, in the Erasmus MC - University Medical Center, a 'pre-patient' training curriculum has been implemented, containing a simulator training course until a plateau phase is reached. In our study, this occurred around 60 VR colonoscopies performed on the GI Mentor.

CONCLUSIONS

Virtual reality training on a simulator in a 'pre-patient' training curriculum leads to a significant improvement of performance on the simulator itself and, more importantly, to significantly improved performances during patient-based colonoscopy. This study demonstrates the rationale for intensive simulator training in the early learning curve of novices performing colonoscopy and provides a practical approach to define the extent of VR training that is useful.

REFERENCES

1. Van Sickle KR, Buck L, Willis R, et al. A multicenter, simulation-based skills training collaborative using shared GI mentor II systems: Results from the Texas association of surgical skills laboratories (TASSL) flexible endoscopy curriculum. Surg Endosc Interv Tech 2011;25:2980-2986
2. Ende A, Zopf Y, Konturek P, et al. Strategies for training in diagnostic upper endoscopy: A prospective, randomized trial. Gastrointest Endosc 2012;75:254-260
3. Ahad S, Boehler M, Schwind CJ, Hassan I. The effect of model fidelity on colonoscopic skills acquisition. A randomized controlled study. J Surg Educ 2013;70:522-527
4. Ahlberg G, Hultcrantz R, Jaramillo E, Lindblom A, Arvidsson D. Virtual reality colonoscopy simulation: a compulsory practice for the future colonoscopist? Endoscopy 2005;37:1198-1204
5. Baillie J, Evangelou H, Jowell P, Cotton PB. The future of endoscopy simulation: a Duke perspective. Endoscopy 1992;24 Suppl 2:542-543
6. Gerson LB, Van Dam J. Technology review: the use of simulators for training in GI endoscopy. Gastrointest Endosc 2004;60:992-1001
7. Lim BS, Leung JW, Lee J, et al. Effect of ERCP mechanical simulator (EMS) practice on trainees ERCP performance in the early learning period: US multicenter randomized controlled trial. Am J Gastroenterol 2011;106:300-306
8. Koch AD, Buzink SN, Heemskerk J, et al. Expert and construct validity of the Simbionix GI Mentor II endoscopy simulator for colonoscopy. Surg Endosc 2007
9. Koch AD, Haringsma J, Schoon EJ, de Man RA, Kuipers EJ. A second-generation virtual reality simulator for colonoscopy: validation and initial experience. Endoscopy 2008;40:735-738
10. Frank JR, Langer B. Collaboration, communication, management, and advocacy: teaching surgeons new skills through the CanMEDS Project. World J Surg 2003;27:972-978; discussion 978
11. 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
12. Sedlack RE, Kolars JC, Alexander JA. Computer simulation training enhances patient comfort during endoscopy. Clin Gastroenterol Hepatol 2004;2:348-352
13. Grantcharov TP, Reznick RK. Teaching procedural skills. Bmj 2008;336:1129-1131
14. Walsh CM, Sherlock ME, Ling SC, Carnahan H. Virtual reality simulation training for health professions trainees in gastrointestinal endoscopy. Cochrane Database Syst Rev 2012;6:CD008237
15. McConnell RA, Kim S, Ahmad NA, et al. Poor discriminatory function for endoscopic skills on a computer-based simulator. Gastrointest Endosc 2012
16. Sugden C, Aggarwal R, Banerjee A, et al. The development of a virtual reality training curriculum for colonoscopy. Ann Surg 2012;256:188-192
17. Felsher JJ, Olesevich M, Farres H, et al. Validation of a flexible endoscopy simulator. Am J Surg 2005; 189:497-500
18. Gillies D, Haritsis A, Williams C. Computer simulation for teaching endoscopic procedures. Endoscopy 1992;24 Suppl 2:544-548
19. Mahmood T, Darzi A. A study to validate the colonoscopy simulator. Surg Endosc 2003;17:1583-1589
20. Sedlack RE, Kolars JC. Validation of a computer-based colonoscopy simulator. Gastrointest Endosc 2003;57:214-218
21. Buzink SN, Koch AD, Heemskerk J, et al. Acquiring basic endoscopy skills by training on the GI Mentor II. Surg Endosc 2007;21:1996-2003
22. Mahmood T, Darzi A. The learning curve for a colonoscopy simulator in the absence of any feedback: no feedback, no learning. Surg Endosc 2004;18:1224-1230
23. Verdaasdonk EG, Stassen LP, van Wijk RP, Dankelman J. The influence of different training schedules on the learning of psychomotor skills for endoscopic surgery. Surg Endosc 2007;21:214-219

24. Eversbusch A, Grantcharov TP. Learning curves and impact of psychomotor training on performance in simulated colonoscopy: A randomized trial using a virtual reality endoscopy trainer. Surg Endosc 2004;18:1514-1518
25. Sedlack RE, Kolars JC. Computer simulator training enhances the competency of gastroenterology fellows at colonoscopy: Results of a pilot study. Am J Gastroenterol 2004;99:33-37
26. Buzink SN, Koch AD, Heemskerk J, et al. Acquiring basic endoscopy skills by training on the GI Mentor II. Surg Endosc Interv Tech 2007;21:1996-2003
27. Ahlberg G, Hultcrantz R, Jaramillo E, Lindblom A, Arvidsson D. Virtual reality colonoscopy simulation: A compulsory practice for the future colonoscopist? Endoscopy 2005;37:1198-1204
28. Ferlitsch A, Schoefl R, Puespoek A, et al. Effect of virtual endoscopy simulator training on performance of upper gastrointestinal endoscopy in patients: A randomized controlled trial. Endoscopy 2010;42:1049-1056
29. Cohen J, Cohen SA, Vora KC, et al. Multicenter, randomized, controlled trial of virtual-reality simulator training in acquisition of competency in colonoscopy. Gastrointest Endosc 2006;64:361-368
30. 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
31. 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 2009

Chapter 8

Competence measurement during colonoscopy training: the use of self-assessment of performance measures

Arjun D. Koch[1], Jelle Haringsma[1*], Erik J. Schoon[2], Rob A. de Man[1], Ernst J. Kuipers[1]

1)Department of Gastroenterology and Hepatology, Erasmus MC - University Medical Center Rotterdam, the Netherlands

2)Department of Gastroenterology and Hepatology, Catharina Hospital Eindhoven, the Netherlands

**) Current affiliation: Department of Gastroenterology and Hepatology, Maasstad Hospital Rotterdam, the Netherlands*

Am J Gastroenterol 2012;107:971–975

ABSTRACT

introduction

Colonoscopy is a practical skill requiring extensive training. To date there has been little attempt to comprehensively assess both generic and specific technical skills in flexible endoscopy. Our aim was to evaluate a newly developed assessment technique for colonoscopy training.

methods

We prospectively evaluated colonoscopy skills during training using the Rotterdam Assessment Form for colonoscopy (RAF-c). The questionnaire covers objective data on cecal intubation, procedural time and subjective grading of performance. The individual learning curves in colonoscopy are compared to a group reference.

results

A total of 2887 colonoscopies were self-assessed by 19 trainees. The cecal intubation rate improved from 65% (range 45-80%) at baseline, to 78% (range 65-95%) and 85% (range 80-90%) after performing 100 and 200 colonoscopies respectively. In our training program the 90% threshold was reached after an average of 280 colonoscopies. Cecal intubation time improved from 13:10 minutes (range 9-19) at baseline, to 9:30 (range 7-13) and 8:30 (range 7-10) after performing 100 and 200 colonoscopies respectively.

conclusions

This novel self-assessment form provides insight in individual learning curves of colonoscopists in training. Individual learning curves can be compared to a group reference. It provides data on the development of dexterity skills as well as individual training targets, and stimulates trainees to identify steps for self-improvement. Procedural competence, i.e. 90% cecal intubation, was reached at an average 280 colonoscopies.

INTRODUCTION

Colonoscopy requires practical skills obtained by extensive training. Colonoscopy dexterity has to be acquired over several hundred procedures. Trainees differ considerably in the rate at which they acquire the necessary psychomotor skills, although evidence suggests that most will arrive at a common endpoint, named procedural competence.

Measures for colonoscopy competence are however ill defined. Only a few studies have been published on colonoscopy competence acquisition[1-4]. These studies particularly focused on procedure numbers as a pseudo-marker for competence. Recently, studies have been published on novel assessment tools for procedural competency in colonoscopy and learning curves[5-8]. Both the DOPS (Direct Observation of Procedural Skills) and the MCSAT (Mayo Colonoscopy Skills Assessment Tool) have been published as validated tools to assess procedural skills.

For total colonoscopy, cecal intubation is a prerequisite. Cecal intubation rate is therefore often used as a measure of procedural competence. Generally, 90% cecal intubation rate is considered the minimum threshold for competence[1-4, 9]. However, cecal intubation rate as a primary endpoint provides no insight in the manner that the cecum is reached, in how much time, and with which effort and patient burden.

The current guidelines[10-12] by the Accreditation Council for Graduate Medical Education (ACGME) and the American Society for Gastrointestinal Endoscopy (ASGE) state that procedural competence should be based on objective performance criteria. However, the only criterion which these guidelines define is the number of procedures performed. As such, endoscopists in the US can be certified for colonoscopy after the minimum threshold number of 140 procedures. This threshold number is based on a study by Cass et al.[1]. In this study, the authors described that 35 trainees reached a cecal intubation rate of 90% after having performed 140 colonoscopies. In 1999, the same authors reported the combined results of 11 studies on cecal intubation rates, showing that the 90% threshold was reached after 349 colonoscopies[2]. This number was much more in concordance with a previously published prospective multicenter study[4].

Few attempts have been made to assess individual performance progress over time using group performance as a reference. In colonoscopy, such appraisal could be used for certification, but also to gain insight in the factors that determine the learning process. Furthermore, by comparing individual growth to a group reference, deviation can be noticed at an early stage and provide the opportunity to intervene if necessary. In this article we describe a simple method that can provide such data.

METHODS

From January 2008 until December 2010 we performed a prospective, single center evaluation of individual and group colonoscopy performance progress in colonoscopy trainees. In the Netherlands, training for gastroenterology starts with 2 year training in Internal Medicine, followed by 4 year training in Gastroenterology and Hepatology. All fellows perform endoscopy throughout this 4 year training period. All residents in our training program were included. Immediately after each colonoscopy procedure, trainees completed a simple, newly developed self-assessment form. The self-assessment method was performed on top of the routine master-apprentice training program for colonoscopy.

participants

In our institution all trainees enter the colonoscopy training program after a standardized tutorial on the fundamentals of the endoscopy kit and colonoscopy technique. All trainees completed a two-day training to ensure basic colonoscope handling. This program provides intensive training using endoscopy simulators (Simbionix GI Mentor II, Simbionix, Tel Hashomer, Israel, and Olympus Endo TS-1, Olympus KeyMed, Southend, UK), with continuous instruction and feed-back by dedicated teachers.

self-assessment form

For this program, the Rotterdam Assessment Form for colonoscopy (RAF-c) was used. The form in it self is an unvalidated assessment tool; however, it is partially based on previously validated assessment tools[6, 13-14]. The form consists of three parts (Figure 1). The first part scores objective parameters such as cecal intubation and the time to reach the cecum. In the second part, the trainee is asked to self-evaluate performance on six aspects using visual analogue scales. After every ten procedures the trainee is asked to formulate a short-term learning plan, based on the colonoscopy performance at that stage. This is done using a universal four-step approach to improvement used by teachers in various fields, based on the Osborn-Parnes Creative Problem Solving Process developed in the 1950s[15].

patients

All procedures and report writing were performed as part of the regular training program, supervised by a staff endoscopist. Patients were referred for colonoscopy for generally accepted indications or for colorectal cancer screening. Demographic data were collected on age, gender, previous colonoscopies and previous abdominal surgery. Our patient population was comparable to previous studies on the subject. Routinely, patients are sedated using fentanyl and midazolam, however at the request of the individual patient this management is sometimes changed. There was no comparison of cecal intubation rates in sedated and unsedated procedures. The study was approved by the Institutional Review Board of the Eras-

1. **Objective assessment:**

Reached the cecum? ☐ yes → What time? ………. min
☐ no

1.2 **failed to reach cecum:**
Reached Sigmoid? ☐ yes → Time? ………. min
☐ no

Reached Flexure? ☐ yes → ☐ splenic ☐ hepatic Time? ………. min
☐ no

1.3 Time spent colonoscopy independently? ………. Min

2. **Subjective assessment:**

How do you rate the colonoscopy procedure you just performed:

Visual Analogue Scale

0 (Very bad) (Very good) **10**

2.1 Time to reach cecum
2.2 Colonoscope handling
2.3 Loopings solved
2.4 Time spent with clear view
2.5 Safety of procedure
2.6 Overall performance

3. **Improvement plan:** (Define potential points for improvement)

What is the situation?______________________________

What is the problem?______________________________

How should it be addressed?__________________________

What is the improvement strategy?_______________________

Figure 1. The Rotterdam Assessment Form for colonoscopy (RAF-c). A self-assessment form for colonoscopy

mus Medical Center. Patients provided informed consent for trainee endoscopy and evaluation of colonoscopy performance as specified.

outcome measures

The primary outcome measure was cecal intubation. Cecal intubation was documented by identification and photo-documentation of the appendiceal orifice and the ileo-cecal valve as in previous studies[9, 16]. The time interval from insertion to cecal intubation was used as a secondary outcome measure. Successful colonoscopy by the trainee was defined as cecal intubation without any physical assistance by the supervisor. If the supervising endoscopist had to take over for any reason, the procedure was considered failure for the trainee. Supervisors were present during all colonoscopies performed by trainees. Verbal assistance was given whenever the staff member felt necessary or when asked by the trainee. All procedures were documented in the electronic patient record system of the hospital, which has full coverage of all procedures (Endobase®, Olympus-Europe, Hamburg, Germany). Reports had to be authorized by the supervising staff member. Individual adherence to completion of the RAF-c was calculated comparing the number of completed assessment forms to procedures performed as extracted from the report system.

statistical analysis

Group averages and standard deviations were calculated using SPSS 15.0. One-way ANOVA was used to calculate significance in increase in the group learning curve. A two-side p-value of < 0.05 was considered significant. Graphs and trend lines were created using standard software. Graphs were created using moving averages for the percentage of cecal intubation.

RESULTS

From January 2008 until December 2010, 19 trainees performed a total of 2887 colonoscopies, corresponding with 152 procedures per trainee (range 91-347). The mean patient age was 49.6 years, 48% male and 52% female. 30% of patients had undergone previous colonoscopies for various reasons and 22% of patients had a history of abdominal surgery. Thirteen residents were novices in colonoscopy and were assessed form the start of their training. Six were senior residents with an average number of 204 colonoscopies performed before the start of the study.

Initial adherence to RAF-c completion varied from 45% to 100% after 20 colonoscopies. This increased to an average of 91% (range 84-100%) after subsequent feedback. Analysis of the procedures that were skipped for assessment by the trainees showed no significant difference in cecal intubation rate compared to those that were assessed. The median time to complete the RAF-c was 36 seconds (range 20-120 seconds). The group average for cecal in-

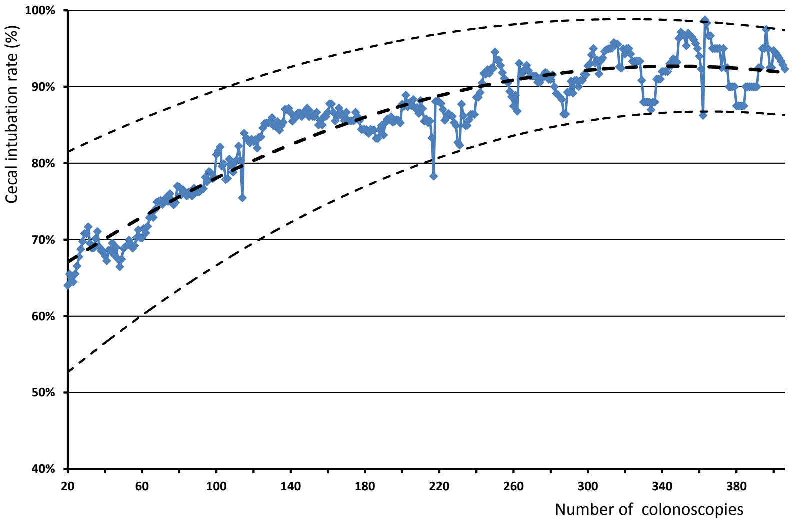

Figure 2. Group reference on cecal intubation rate with standard deviations
Blue markers: group average
Black trendlines: group average (bold) and standard deviations +1SD and -1SD (normal)

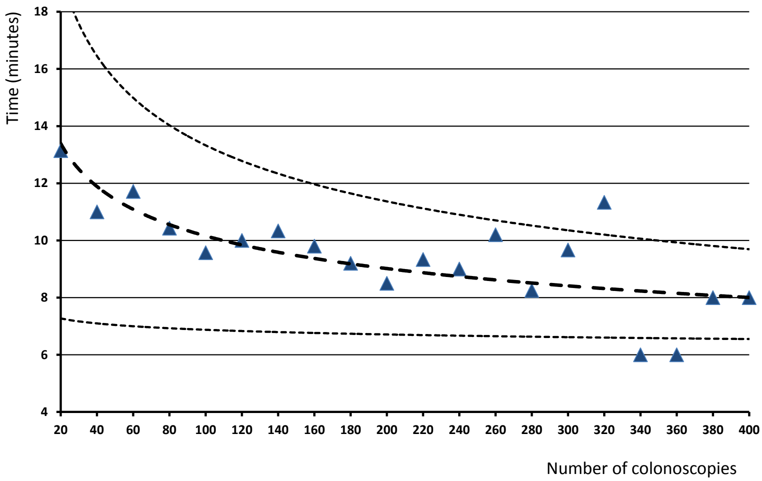

Figure 3. Group reference on cecal intubation time with standard deviations
Blue markers: group average
Black trendlines: group average (bold) and standard deviations +1SD and -1SD (normal)

tubation and time to reach the cecum are shown in Figures 2 and 3. Baseline cecal intubation in the first 20 procedures averaged 65% (range 45-80%). At 100 procedures, cecal intubation rose to 78% (range 65-95%), at 200 procedures to 85% (range 80-90%). The 90% threshold for cecal intubation was crossed at a mean of 280 colonoscopies (range 144-346 procedures). Overall progress of the group learning curve was statistically significant with $P < 0.001$ using one-way ANOVA.

Cecal intubation time was 13:10 minutes (range 9-19) at baseline. After 100 colonoscopies, cecal intubation time decreased to 9:30 (range 7-13), after 200 procedures to 8:30 (range 7-10).

A deviating learning curve of one of our trainees is shown in Figure 4. This trainee had a very slow start. The problem was mostly contributed to loop management. This information was extracted from her self-assessments as well as interpretation by the supervising endoscopists. The problem was managed by providing extra theoretical background on loop handling, simulator training and ScopeGuide assessment during her colonoscopies (ScopeGuide™ Olympus, Tokyo, Japan). These interventions led her back on track performing on the group average.

An improvement plan was filled out in 36% of procedures (range 10-70%). In 184 of 288 cases there was no improvement plan defined. Most plans dealt with recognition and solving of loops, straightening the endoscope and slowing down in difficult situations. Most trainees concluded they needed to perform more procedures in order to master the required skills.

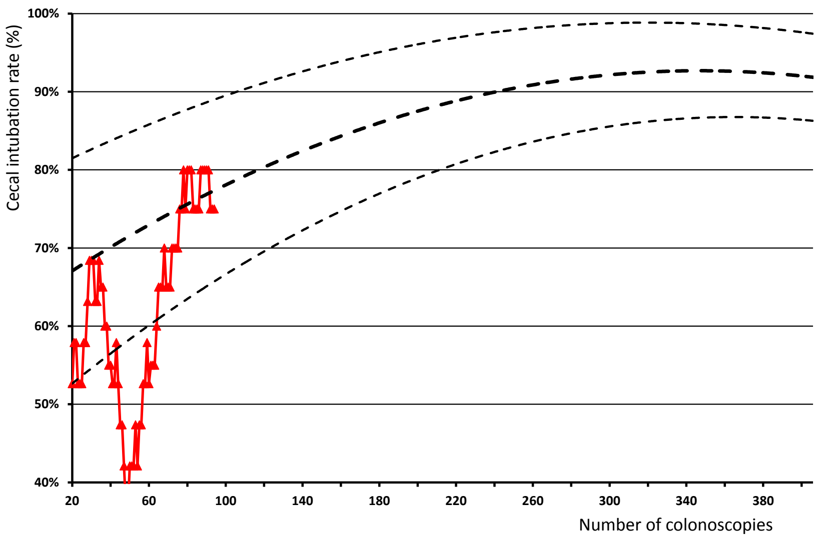

Figure 4. Example of a learning curve of an individual trainee performing under average
Red markers: individual performance
Black trendlines: group average (bold) and standard deviations +1SD and -1SD (normal)
This trainee actually had a very slow start. The problem was mostly contributed to loop management. This information was extracted from her self-assessments as well as interpretation by the supervising staff members. The problem was managed by providing her with extra theoretical background on loop handling, simulator training and ScopeGuide assessment during her colonoscopies. These interventions led her back on track performing on the group average.

More detailed plans included using the ScopeGuide imager, training specific loops on the simulators or ask for more guidance by the supervising endoscopist.

DISCUSSION

Colonoscopy is the standard technique for diagnosis and treatment of ileo-colonic disorders. It plays a pivotal role in colorectal cancer screening and prevention. This explains that colonoscopy is performed in approximately 14.2 million cases in the US annually[17]. Colonoscopy can be burdensome, with a marked risk for complications[18]. Advancing the endoscope is technically demanding, with a high variety in quality and success[19]. Recognized markers for quality are cecal intubation rates, withdrawal times[20], adenoma detection rates and colon cancer miss rates[21].

Only a few studies on colonoscopy skills acquisition have been performed. Most of these focused on threshold numbers necessary for certification. There were marked differences in outcome between studies with numbers of procedures needed for cecal intubation success ranging from 100 to 1048 procedures. There was no assessment of individual performance against group learning curves. By monitoring a group of trainees from the level of novices to proficient endoscopists, we have developed a method to monitor learning that can be easily incorporated into daily practice. Based on individual data, a group learning curve can be built for a training program. This learning curve can be used as a group reference in order to assess individual performance. Hence, procedural competency can simply be benchmarked against objective standards instead of a non-validated threshold number of procedures.

With respect to average numbers of procedures needed to reach the threshold cecal intubation rate, our results seem comparable to previous studies[1-4, 9]. Recently, similar results from the Mayo clinic were published[8]. In this study an 85% cecal intubation rate was achieved after 250 to 275 procedures and this percentage increased to 92% after 300 procedures.

No recognized threshold exists for cecal intubation time. This secondary parameter is merely used to validate that the improvement in cecal intubation over time reflects enhanced skill. The assessment of multiple contrary dexterity parameters has been proven useful. During skills assessment and training of experienced colonoscopists, Valori et al. compared cecal intubation time to patient pain levels. This made endoscopists aware of the fact that a higher cecal intubation rate "at all costs" might compromise patient safety[5, 22].

Our self-evaluation method is used as an add-on to the traditional master-apprentice model. This evaluation is biased and subject to a trainee's knowledge. Yet, the evaluation is likely to be influenced by input from the trainer. Furthermore, self-maintenance is also more likely to raise self-awareness and a critical self-attitude. This is in agreement with data from other learning programs. Together, we believe this provides the add-on effect of self-evaluation in optimizing the learning process.

In concordance with the ACGME and ASGE, we believe that in the current era of training, threshold numbers of procedures should be replaced by learning curves to monitor competence performance[10-12]. Learning curves are a valuable means to assess competence when compared to a group reference. The Rotterdam Assessment Form for colonoscopy is a rapid and easy tool to monitor these learning curves. Secondly, it provides a platform to create self-awareness in learning objectives in endoscopy. The feedback mechanism in the improvement plan was severely under-utilized. In the future we plan to increase adherence by stimulating trainees that perform under average. In the Netherlands, the Rotterdam Assessment Form for colonoscopy is incorporated in a newly developed e-portfolio available for all gastroenterologists in training.

REFERENCES

1. Cass O, Freeman ML, Cohen J *et al.* Acquisition of Competency in Endoscopic Skills (ACES) during Training: a Multicenter Study. Gastrointest Endosc 1996;43:308.
2. Cass OW. Training to competence in gastrointestinal endoscopy: a plea for continuous measuring of objective end points. Endoscopy 1999;31:751-4.
3. Freeman ML. Training and competence in gastrointestinal endoscopy. Rev Gastroenterol Disord 2001;1:73-86.
4. Marshall JB. Technical proficiency of trainees performing colonoscopy: a learning curve. Gastrointest Endosc 1995;42:287-91.
5. Gerson LB, Van Dam J. Technology review: the use of simulators for training in GI endoscopy. Gastrointest Endosc 2004;60:992-1001.
6. Colonoscopy. DOPS assessment form. The Joint Advisory Group on GI Endoscopy. Volume 2010, 2008. http//:www.thejag.org.uk/Forms/tabid/72/Default.aspx Accessed [September 5, 2010]
7. Sedlack RE. The Mayo Colonoscopy Skills Assessment Tool: validation of a unique instrument to assess colonoscopy skills in trainees. Gastrointest Endosc 2010;72:1125-33
8. Sedlack RE. Training to competency in colonoscopy: assessing and defining competency standards. Gastrointest Endosc 2011;74(2):355-366
9. Thomas-Gibson S. The caecum or not the caecum? Eur J Gastroenterol Hepatol 2008;20:500-2.
10. ACGME. ACGME Program Requirements for Graduate Medical Education in Gastroenterology, 2007.
11. Wexner SD, Litwin D, Cohen J *et al.* Principles of privileging and credentialing for endoscopy and colonoscopy. Gastrointest Endosc 2002;55:145-8.
12. ASGE/ACG. ASGE/ACG Taskforce on Quality in Endoscopy. Ensuring Competence in Endoscopy. Available at: http://www.asge.org/WorkArea/showcontent.aspx?id 3384. Accessed May 30, 2011.
13. Park J, MacRae H, Musselman LJ *et al.* Randomized controlled trial of virtual reality simulator training: transfer to live patients. Am J Surg 2007;194:205-11.
14. Reznick R, Regehr G, MacRae H *et al.* Testing technical skill via an innovative "bench station" examination. Am J Surg 1997;173:226-30.
15. Francis NK, Hanna GB, Cresswell AB *et al.* The performance of master surgeons on standard aptitude testing. Am J Surg 2001;182:30-3.
16. Rex DK. Still photography versus videotaping for documentation of cecal intubation: a prospective study. Gastrointest Endosc 2000;51:451-9.
17. Seeff LC, Richards TB, Shapiro JA *et al.* How many endoscopies are performed for colorectal cancer screening? Results from CDC's survey of endoscopic capacity. Gastroenterology 2004;127:1670-7.
18. Panteris V, Haringsma J, Kuipers EJ. Colonoscopy perforation rate, mechanisms and outcome: from diagnostic to therapeutic colonoscopy. Endoscopy 2009;41:941-51.
19. 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;97:1296-308.
20. Barclay RL, Vicari JJ, Doughty AS *et al.* Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. N Engl J Med 2006;355:2533-41.
21. Kaminski MF, Regula J, Kraszewska E *et al.* Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med 2010;362:1795-803.
22. Jakimowicz JJ, Cuschieri A. Time for evidence-based minimal access surgery training--simulate or sink. Surg Endosc 2005;19:1521-2.

Chapter 9

General discussion and conclusions

In the first section of this thesis, we presented a systematic review of all available evidence on training and competence assessment in GI endoscopy. A striking conclusion is that after years of research, we are still on the verge of grasping the complexity of training novices in GI endoscopy, assessing their competence level in an objective and universally reproducible fashion with establishment of benchmarks for training and certification. During recent years, the focus of attention has gradually shifted to the concept of a more competence-based training approach instead of the use of threshold numbers. However, with this growing awareness, it has become clear that validated tools to measure performance are scarce. In the studies that were carried out to form the basis of this thesis, we have tried to gain more insight in the acquisition of skills necessary to perform colonoscopy and the role that simulators can play in training novice colonoscopists.

Virtual Reality (VR) endoscopy simulators offer a promising tool for teaching basic dexterity prior to training in patient-based colonoscopy. Training can be given in a learning environment, without putting patients at unnecessary discomfort or risk.[1-3] This allows for trainees to train their skills and repeat exercises until specific endoscopy skills are mastered. Furthermore it enables trainees to test the boundaries of certain maneuvers like insufflation or endoscope pushing with regards to the effect on discomfort of the virtual patient. Before implementing the use of novel simulators into training programs, it is essential to carry out studies to determine face and construct validity.[4] Only such a study can demonstrate the validity and reliability of using a simulator as a training or assessment tool. Validation forms the basis on which further studies can be conducted.

In this thesis we have demonstrated both face and construct validity for the Simbionix GI Mentor II™ and Olympus Endo TS-1™ for colonoscopy. This means that both machines simulate colonoscopy to a favorable degree of realism according to expert endoscopists and that the simulators are able to distinguish competency levels of different users with different experience levels. This enables the simulator to notice improvement of novice trainees after repeated exercises. A possible pitfall of repeated exercises is the learning effect of the exercise itself instead of improvement of skills. By having expert endoscopists repeating the same exercises over and over, we demonstrated that they did not reach higher performance levels on the simulator. This means that the test-retest reproducibility or internal consistency of the simulators is highly reliable and improved performance on the simulator does indeed reflect improved skills. The construct validity is mostly based on cecal intubation rates and procedural times. This is similar for all endoscopy simulators that are currently commercialized. The machines still lack the power to demonstrate differences in other performance outcome metrics like simulated pain or unsafe maneuvers.

We trained 30 novices during four sessions in which they all performed 15 VR simulator tasks on the GI Mentor II. All participants improved significantly in their time to cecal intubation and VR dexterity game, but did not yet reach the level of experienced endoscopists. We predicted that this level of performance would be reached between 50 and 100 procedures.

We later confirmed this number to be around 60 procedures during prolonged training on the simulator and transfer of skills to patient-based colonoscopy.

With the knowledge that repeated training improves performance, we designed a study to explore how much simulator-based training is useful and how this impacts patient-based colonoscopy performance. Two separate cohorts of novice colonoscopists followed a training curriculum of either 50 or 100 simulator-based colonoscopies. Patient-based assessment was carried out after 10, 30 and 50 simulator colonoscopies for the first group and after 20, 60 and 100 colonoscopies for the second group. This method was chosen to minimize the learning effect of each patient-based assessment. Because of the limited experience of the participants, every extra procedure was considered to possibly have a major impact on their competence level. We demonstrated significant improvement of performance by training on the GI Mentor II up to around 50 colonoscopies. There was no further improvement after 60 simulator procedures, neither during virtual, simulator endoscopy, nor during patient-endoscopy. We conclude that novices gain significant experience by training on simulators before they are exposed to patients. We can also conclude that the added value of simulator training on patient-based performance ceases when the learning curve on the simulator levels out. For the GI Mentor II, this pivotal point appears to average around 60 procedures.

A different approach was chosen in investigating the transfer of skills to patient-based colonoscopy using the Endo TS-1 as a training tool compared to bedside training. 36 Participants were randomly allocated to either simulator-based or patient-based colonoscopy training for 16 hours. Assessment of both groups during patient-based colonoscopy demonstrated similar results. The simulator-trained subjects outperformed the controls during simulator-based assessment. This might be explained by a test-retest reproducibility error but can also very well be explained by the fact that the simulator-trained group received a very structured training on the simulator, dealing with resolution of loops and problems for all colonic segments. The control group was only able to learn from situations that actually occurred during their patient-based training.

In the final section of this thesis we aimed to gain insight in the learning curve during patient-based colonoscopy to the level of procedural competency. A single center prospective registry was carried out using the Rotterdam Assessment Form for Colonoscopy, or RAF-C, as an assessment tool. In this study 19 trainees self-assessed almost 3000 procedures. We found that on average the level of procedural competency measured by means of cecal intubation rate and timing was reached around 280 colonoscopies. The learning curve demonstrated the progress over time and was unmistakably able to demonstrate whenever a steady competent level was reached. The group reference was used to identify individual trainees who were not performing according to the average at any point in time. This allowed us to be able to intervene whenever a trainee was performing under average or when increment diminished, compared to what was expected. The unique improvement plan, which was built into the form, allowed for self-assessment by the trainee or by the trainer to evaluate the problem and come up with possible solutions and steps to correct the omissions at that specific learning point. This improvement plan is a universal four-step approach to improvement used by teachers in various fields, based on the Osborn–Parnes Creative Problem Solving Process developed in the 1950s. It shows large similarities to the Plan-Do-Check-Act cycle of Deming and the Kolb experiential learning cycle. As it turned out, the self-improvement plan was severely under-used during our study by most participants. However, it undoubtedly showed its value for one trainee who was performing under average. The learning objectives, that were generated through her improvement plans, together with additional structured training based on these objectives, lead her back on track performing according to the group average in very little time.

The use of learning curves to monitor progression towards a procedural competency endpoint offers much more potential. Currently we have only explored the learning curves for cecal intubation. In our opinion colonoscopy in practice or training can be divided in three domains. The first domain (I) is formed around cecal intubation, the prerequisite for a complete inspection of the entire colon. This domain includes the acquisition of the necessary skills to handle loop formation and deal with patient comfort. The second domain (II) consists of adequate and complete mucosal inspection and recognition of pathology. The third domain (III) consists of endoscopic therapy such as complete and safe polypectomy. During training every endoscopist has to develop the necessary skills in each of these domains to become a competent and certified endoscopist. This means that the ideal assessment form should at least cover these 3 domains. It is however not appropriate to incorporate an adenoma or polyp detection rate (ADR or PDR) as a measure for competency development during training. ADR and PDR are highly valuable quality parameters, but are typically measured over several hundreds or thousands of procedures. The learning curve of novices typically demonstrates a steep increase in performance up to the level of procedural competence within 300 procedures. Significant improvement can be observed within blocks of 20 sequential colonoscopies. The ADR and PDR lack the discriminative power on such small numbers. Currently, the best way to assess the second and third domain is by direct observation by a supervising expert. Finally, by identifying the fast and slow learners within a group of trainees, fast learners can proceed earlier to advanced programs, allowing slow learners to complete more procedures and thus providing optimal allocation of training capacity.

IMPLEMENTATION

Our current training program for flexible endoscopy has undergone rigorous changes. Virtual reality endoscopy simulator training is implemented in a two-day structured, hands-on training program which takes place at the Erasmus MC Skillslab, organized by the Erasmus School of Endoscopy (ESE). This course is obligatory for all novices entering the gastroenterology training program at the Erasmus MC – University Medical Center. The next logical step would be to provide a prolonged simulator training program to harvest the maximum effect of simulator-based training. This step is not yet in place.

Trainees proceed to patient-based endoscopy with continuous self-assessment and periodical supervised assessment. This is all documented in a nationwide e-portfolio. Trainees progress to more advanced endoscopy programs after repeated proven procedural competence levels.

FUTURE PERSPECTIVES

Efforts should be made in order to continuously improve endoscopy simulators. Although they have undoubtedly proven their value, endoscopy simulators still lack the degree of realism compared to for instance flight simulators. We are convinced that, a closer resemblance to real colonoscopy will lead to increased skills acquisition during simulator-based training and decreased patient burden during the early learning curve. Another role for simulator training might be in training complex loop-formations and resolutions for more experienced endoscopists. This prospective has not been explored.

Simulator-based competency assessment is a subject worthy of investigation. It offers the potential for a highly objective assessment in the most reproducible way. So far, efforts have been disappointing but it is a topic that is currently under investigation.

We believe that we have described a novel method to track competency development for colonoscopy and cecal intubation. The next steps should focus on colonoscopy withdrawal techniques and routine intervention techniques like training polypectomy. Competency assessment tools should be available for all three domains within the field of diagnostic and therapeutic endoscopy.

Similar studies are currently undertaken to explore learning curves and assessment tools for ERCP. Endosonography should be the next step. Especially since, during the past decade, endosonography has transformed from not just a diagnostic tool, to a high level therapeutic procedure.

REFERENCES

1. Sedlack, R.E., J.C. Kolars, and J.A. Alexander, *Computer simulation training enhances patient comfort during endoscopy*. Clin Gastroenterol Hepatol, 2004. 2(4): p. 348-52.
2. Thomas-Gibson, S. and C.B. Williams, *Colonoscopy training--new approaches, old problems.* Gastrointest Endosc Clin N Am, 2005. 15(4): p. 813-27.
3. Grantcharov, T.P. and R.K. Reznick, *Teaching procedural skills.* Bmj, 2008. 336(7653): p. 1129-31.
4. Carter, F.J., et al., *Consensus guidelines for validation of virtual reality surgical simulators.* Surg Endosc, 2005. 19(12): p. 1523-1532.

NEDERLANDSE SAMENVATTING

In het eerste deel van dit proefschrift hebben we een systematische review verricht van alle beschikbare gegevens over opleiding en competentie beoordeling in gastrointestinale (GI) endoscopie. Een opvallende conclusie is dat we na jaren van onderzoek nog altijd maar aan het begin staan van het doorgronden van de complexiteit van de opleiding in GI endoscopie en de beoordeling van een het competentieniveau op een objectieve en universeel reproduceerbare wijze volgens vastgestelde criteria voor opleiding en certificering. In de afgelopen jaren is het focus voor opleiding geleidelijk verschoven van een meer competentiegerichte aanpak van de opleiding in plaats van het gebruik van tevoren vastgestelde aantallen verrichtingen. Echter, hiermee is het pijnlijk duidelijk geworden dat gevalideerde instrumenten om prestaties te meten schaars of niet voorhanden zijn. In de studies die wij hebben uitgevoerd en de basis van dit proefschrift vormen, hebben we getracht om meer inzicht te krijgen in de verwerving van vaardigheden die nodig zijn om colonoscopie uit te voeren. Daarnaast hebben we onderzocht welke rol simulatoren kunnen spelen in de opleiding van beginnende colonoscopisten.

Virtual Reality (VR) endoscopie simulatoren vormen een veelbelovend platform voor het aanleren van basale vaardigheden voor de training van colonoscopie voorafgaand aan de praktijk. Training kan op deze manier worden verzorgd in een leeromgeving, zonder onnodig ongemak of risico voor patiënten. Het maakt mogelijk dat beginners hun vaardigheden onuitputtelijk kunnen trainen en oefeningen kunnen herhalen tot specifieke endoscopie vaardigheden zijn verworven en worden beheerst. Daarnaast is het in deze leeromgeving mogelijk om de grenzen van bepaalde manoeuvres zoals teveel lucht inblazen of het te ver laten uitbochten van de endoscoop duwen te ervaren met betrekking tot het effect op ongemak van de virtuele patiënt. Het is van essentieel belang dat deze simulatoren uitvoerig onderzocht worden om te bepalen in welke mate de simulatoren inderdaad de werkelijkheid nabootsen voordat ze worden ingezet in trainingsprogramma's. Dit houdt in dat de validiteit getest moet worden. Een van de manieren om dit te doen is door de zogenaamde face en construct validity te testen. Deze validatie vormt de basis waarop verdere studies kunnen worden uitgevoerd.

In dit proefschrift hebben we aangetoond dat zowel de Simbionix GI Mentor II™ en Olympus Endo TS-1™ valide instrumenten zijn voor de nabootsing van colonoscopie. Dit betekent dat beide machines colonoscopie simuleren die in hoge mate overeenkomt met de realiteit en dat de simulatoren in staat zijn om competentie niveaus van verschillende gebruikers met wisselende ervaring van elkaar te onderscheiden. Hierdoor is de simulator in staat om verbetering van beginnende endoscopisten objectief vast te leggen na herhaalde oefeningen. Een mogelijke valkuil van herhaalde oefeningen is het leereffect van de oefening zelf in plaats van het verbeteren van vaardigheid. Wij hebben aangetoond dat als ervaren endoscopisten dezelfde oefeningen herhalen, er geen verbetering van hun prestatie op de simulator optreedt.

Dit betekent dat de test-retest reproduceerbaarheid van de simulatoren zeer betrouwbaar is en verbeterde prestaties op de simulator gemeten inderdaad een afspiegeling zijn van toegenomen vaardigheden. De construct validity is vooral gebaseerd op coecum intubatie en verbetering van de tijd tot coecum intubatie. Dit lijkt het geval te zijn voor alle endoscopie simulatoren die momenteel commercieel verkrijgbaar zijn. Het ontbreekt de simulatoren nog aan onderscheidend vermogen om verschillen in andere uitkomsten zoals gesimuleerde pijn of onveilige manoeuvres aan te tonen.

In een van onze studies trainden wij 30 beginners gedurende vier sessies waarin 15 VR simulator opdrachten op de GI Mentor II werden uitgevoerd. Alle beginners lieten een aanzienlijke verbetering zien in de tijd die nodig was om het coecum te intuberen alsmede ook in een behendigheidsspel. Ze bereikten echter nog niet het niveau van ervaren endoscopisten. Aan de leercurve konden we voorspellen dat dit punt waarschijnlijk zou worden bereikt tussen 50 en 100 procedures.

Met de wetenschap het leereffect van de simulator nog niet ten volle was benut, ontworpen we een studie om te onderzoeken hoeveel simulator training nuttig is en hoe zich dit vertaalt in betere resultaten tijdens colonoscopie in de praktijk. Twee afzonderlijke cohorten van beginnende colonoscopisten werden gevormd. Het eerste cohort verrichtte 50 colonoscopieën op de simulator, het tweede cohort 100 colonoscopieën. Beoordeling van competentie en verkregen vaardigheden van deze simulator training werd uitgevoerd na 10, 30 en 50 scopieën voor de eerste groep en na 20, 60 en 100 scopieën voor de tweede groep door het verrichten van twee colonoscopieën in de praktijk. Het aantal van twee werd gekozen om het leereffect van iedere praktijk colonoscopie te minimaliseren. Vanwege de beperkte ervaring van de deelnemers, werd elke extra procedure beschouwd als van mogelijk grote invloed op hun competentieniveau. Er bleek een significante verbetering van het competentie niveau tot ongeveer 50 colonoscopieën op de simulator en er was geen verdere verbetering meer aantoonbaar na 60 simulator procedures. Deze afvlakking van de leercurve werd zowel in de competentie gemeten door de simulator als tijdens de praktijk verrichte colonoscopie waargenomen. Onze conclusie hieruit is dat beginners veel waardevolle ervaring kunnen opdoen voordat ze worden blootgesteld aan patiënten. We kunnen ook concluderen dat de toegevoegde waarde van verdere training op de simulator niet meer zinvol is als het leereffect op de simulator niet meer meetbaar is. Voor de GI Mentor II lijkt dit punt rond de 60 procedures te liggen.

We hebben een andere aanpak gekozen om het leereffect op de Endo TS-1 te onderzoeken in vergelijking met standaard praktijk training. 36 Deelnemers werden willekeurig ingedeeld voor 16 uur simulator of 16 uur praktijk training. Beoordeling van beide groepen werd opnieuw verricht tijdens colonoscopie in de praktijk. Het resultaat was dat beide groepen een competentieniveau hadden bereikt waar geen verschillen in konden worden aangetoond in de praktijk. De simulator getrainde scopisten presteerden wel beter in het examen op de simulator. Dit kan mogelijk worden verklaard door een test-retest reproduceerbaarheids fout, een leereffect van de oefening zelf, maar kan ook heel goed worden verklaard door het feit dat de simulator getrainde groep een zeer gestructureerde training had gekregen op de simulator waarin stapsgewijs het optreden en oplossen van complexe uitbochtingen in het colon waren getraind. De controlegroep was alleen in staat om te leren van situaties die daadwerkelijk hadden plaatsgevonden tijdens hun praktijk training.

In het laatste deel van dit proefschrift hebben we getracht inzicht te krijgen in de leercurve die ontstaat door training in de praktijk tot het niveau van procedurele competentie. Een prospectieve registratie werd uitgevoerd met behulp van het Rotterdam Assessment Form for Colonoscopy , of RAF-C. Dit is een door onszelf ontwikkeld beoordelingsformulier dat specifiek is ontworpen voor registratie van verrichtingen zoals colonoscopie met de daarbij behorende uitkomsten. In deze studie registreerden 19 endoscopisten in opleiding circa 3000 colonoscopie procedures. We toonden aan dat het gemiddelde niveau van procedurele competentie, gemeten aan de hand van de gemiddelde coecum intubatie uitgezet tegen het aantal uitgevoerde procedures, werd bereikt na circa 280 colonoscopieën. De leercurve laat onmiskenbaar vooruitgang in de tijd zien en definieert wanneer een vooraf gesteld competentie niveau wordt bereikt. Een groepsgemiddelde werd vervolgens gebruikt om individuele endoscopisten te kunnen vervolgen. Hiermee kan op elk moment eenvoudig inzicht verkregen worden wanneer een endoscopist in opleiding zichzelf niet verbetert naar verwachting. Dit inzicht maakt het mogelijk om te kunnen ingrijpen wanneer een scopist onder het gemiddelde presteert of gaat afbuigen van de verwachte leercurve. Het unieke verbeterplan dat aan het formulier is gekoppeld maakt het mogelijk dat of de supervisoren, of de endoscopisten in opleiding zelf, mogelijke problemen signaleren en evalueren om mogelijke oplossingen en maatregelen te genereren op dat specifieke leerpunt. Dit verbeterplan is een universeel vier stappen plan tot verbetering en wordt wereldwijd gebruikt door docenten op verschillende terreinen. Het is gebaseerd op het Osborn-Parnes Creative Problem Solving proces ontwikkeld in de jaren 1950. Het toont grote overeenkomsten met de Plan-Do-Check-Act cyclus van Deming en de Kolb experiential learning cyclus. Het bleek in de studie periode dat het zelf-verbeterplan weinig gebruikt werd door de meeste deelnemers. Echter, het plan bleek zeer waardevol voor een endoscopist in opleiding die duidelijk onder het gemiddelde presteerde. Door het gebruik van het verbeterplan werden duidelijke leerdoelen voor verbetering geformuleerd die leidden tot aanvullende gestructureerde training en in zeer korte tijd tot een terugkeer naar het verwachte competentie niveau van de rest van de groep.

Het gebruik van leercurven om het proces richting procedurele competentie te monitoren biedt veel meer mogelijkheden. We hebben nu alleen maar gekeken naar de leercurve voor coecum intubatie. Colonoscopie in de praktijk of in de opleiding kan worden onderverdeeld in drie domeinen. Het eerste domein (I) wordt gevormd rond het betrouwbaar bereiken van het coecum, de voorwaarde voor een volledige inspectie van de gehele dikke darm. Dit domein omvat de verwerving van de benodigde vaardigheden om bijvoorbeeld uitbochting en lusvorming te herkennen en op te lossen en het minimaliseren van discomfort voor de patiënt. Het tweede domein (II) bestaat uit een adequate en volledige mucosale inspectie en herkenning van pathologie. Het derde domein (III) bestaat uit endoscopische therapie zoals complete en veilige poliep verwijdering. Tijdens de opleiding dient elke endoscopist de noodzakelijke vaardigheden in elk van deze domeinen te verkrijgen om een bevoegd en gecertificeerd endoscopist te worden. Dit betekent dat het ideale beoordelingsformulier het ook in zich heeft om deze 3 domeinen te toetsen. Het lastige is dat kwaliteitsparameters zoals een adenoom of poliep detectie ratio (ADR of PDR) niet geschikt zijn om competentie-ontwikkeling tijdens de training weer te geven. ADR en PDR zijn zeer waardevolle kwaliteit parameters, maar worden doorgaans berekend aan de hand van honderden of duizenden procedures. De leercurve in de opleiding toont zoals we hebben laten zien een sterke toename tot op het niveau van de procedurele competentie binnen 300 procedures. Significante verbetering kan worden waargenomen binnen blokken van 20 opeenvolgende colonoscopieën. De ADR en PDR missen het onderscheidend vermogen om op zulke kleine aantallen verbetering te laten zien. Momenteel is de beste manier om het tweede en derde domein te beoordelen door directe waarneming van een superviserend deskundige. Tot slot, door het identificeren van de snelle en trage leerlingen binnen een groep, is het mogelijk om snelle leerlingen eerder te promoveren naar geavanceerde programma's, waardoor er meer procedures beschikbaar blijven voor trage leerlingen om uiteindelijk hetzelfde competentie niveau te behalen. Dit draagt bij aan een optimalisering van opleidingscapaciteit.

implementatie

Ons huidig trainingsprogramma voor endoscopie heeft rigoureuze veranderingen ondergaan. Simulator training in basale flexibele endoscopie wordt in een tweedaagse gestructureerde cursus uitgevoerd waarin met name veel tijd wordt besteed aan het zelf oefenen. Deze training vindt plaats in het Erasmus MC Skillslab, georganiseerd door de Erasmus School of Endoscopie (ESE). Deze cursus is verplicht voor alle beginnende endoscopisten in de opleiding tot MDL-arts in het Erasmus MC - Universitair Medisch Centrum. Op basis van de bevindingen uit dit proefschrift zou een volgende logische stap een langdurige simulator trainingsprogramma zijn om het maximale effect van een simulator training te bereiken. Deze stap is momenteel nog niet gezet.

Alle endoscopisten in opleiding verrichten continue zelfevaluatie naast directe supervisie en periodieke toetsing. Dit wordt inmiddels allemaal gedocumenteerd in een landelijk e-portfolio. Na het behaald hebben van vooraf gestelde doelen ten aanzien van competentie niveaus, promoveren de endoscopisten in opleiding naar meer geavanceerde endoscopie programma's.

toekomst perspectieven

Er is ruimte voor verbetering aan endoscopie simulatoren tot een hoger niveau van realiteit met uitbreiding van toepassingen. Hoewel de toegevoegde waarde van simulatoren duidelijk is bewezen, bereiken endoscopie simulatoren nog niet het niveau van bijvoorbeeld flight simulators waar piloten in getraind worden. We zijn ervan overtuigd dat, hoe nauwkeuriger de realiteit nagebootst wordt van echte colonoscopie, des te meer zullen de vaardigheden die tijdens simulator training zijn verworven, bijdragen aan toegenomen competentie en hiermee reductie van eventueel discomfort voor patiënten tijdens de vroege leercurve. Simulatoren met een hoge mate van realiteit kunnen potentieel een rol spelen in het trainen van meer ervaren endoscopisten in de herkenning en behandeling van complexe lusvorming. Hier is nog geen onderzoek naar verricht.

Simulatoren bieden in potentie een uitstekend platform voor objectieve beoordeling van competentie. Het biedt de mogelijkheid voor een zeer objectieve beoordeling op de meest reproduceerbare manier door de vastgestelde scenario's. Studies die tot op heden zijn verricht laten nog niet de gewenste resultaten zien. Het is een onderwerp dat momenteel wordt onderzocht.

In dit proefschrift hebben wij een nieuwe methode om competentie ontwikkeling te vervolgen voor colonoscopie. De logische vervolg stappen moeten worden toegespitst op inspectie technieken en therapeutische technieken zoals poliepectomie training. Er moeten evaluatie formulieren ontwikkeld worden om competentie te meten in alle drie de domeinen binnen het gebied van diagnostische en therapeutische endoscopie.

Vergelijkbare studies worden momenteel uitgevoerd om leercurven en evaluatie methoden te ontwikkelen en evalueren voor ERCP. Competentie meting in endosonografie is de volgende stap. Dit is vooral nodig omdat de endosonografie in de laatste decennia een sterke ontwikkelingen heeft doorgemaakt en al lang niet meer een louter diagnostische ingreep is maar tegenwoordig wordt ingezet in een scala aan complexe therapeutische procedures.

DANKWOORD

Met het dankwoord schrijf ik het laatste hoofdstuk aan dit proefschrift. Ik ben blij dat het er dan eindelijk toch is gekomen. Promotie onderzoek doen naast de opleiding tot MDL-arts en tegelijkertijd een academische carrier opbouwen is geen sinecure. Laat nu het grote genieten maar komen…

Allereerst wil ik mijn promotor prof. dr. E.J. Kuipers en copromotoren dr. R.A. de Man en dr. E.J. Schoon bedanken. Beste Ernst en Rob, als de dag van gisteren herinner ik mij mijn sollicitatie voor de opleiding tot MDL-arts in Rotterdam. Het gesprek liep soepel en we hadden gelijk ideeën om de onderzoekslijn die ik in Eindhoven gestart was voort te zetten. Jullie gaven aan de beslissing "op z'n Rotterdams" te nemen. Dat hield in zeer snel doch uiterlijk binnen een week. Met een goed gevoel verliet ik het gesprek om nog even kennis te maken met Jelle en de eventuele mogelijkheden van simulator onderzoek door te nemen. Jullie haalden me weer uit dat gesprek om toch nog wat te bespreken en tot mijn grote verbazing zeiden jullie dat de beslissing al was gevallen: ik was aangenomen voor de opleiding en het promotie onderzoek was een feit! Mijn eerste zin uit volledige verbazing: "dat is wel heel Rotterdams!".

In de jaren erna heb ik genoten van een zeer goede tijd in het Erasmus MC. Ernst, je bent een enorme inspiratie en drijvende kracht waarmee je de afdeling naar grote hoogte en internationale befaamdheid hebt gebracht. Daarnaast heb ik grote bewondering voor hoe jij toch altijd in staat bent om mijn manuscripten om te bouwen waarbij je de essentie zo helder uiteen kunt zetten dat ik het zelf ook beter ga begrijpen terwijl ik dacht dat ik de expert op dat gebied was.

Beste Rob, jij staat voor structuur. Het is werkelijk wonderbaarlijk hoe jij altijd in staat bent om met een helikopter view boven de materie te hangen. Binnen het onderzoek gaf mij dit vaak veel houvast en als ik ergens vastliep dan kon jij vanuit die structuur wel weer de belangrijke lijnen doorzetten. Dezelfde structurele invloed heb je ook over de opleiding gegoten en dat heeft z'n vruchten duidelijk afgeworpen. Ik heb genoten van een zeer goede opleiding en zeer prettige samenwerking.

Beste Erik, sommige mensen hebben op belangrijke kruispunten in je leven een enorme invloed. Voor mij ben jij zo'n persoon. Jou aanstekelijke enthousiasme over maagdarmlever ziekten en onderzoek doen maakte dat ik binnen een week in het Catharina Ziekenhuis de beslissing had gemaakt dat ik MDL-arts wilde worden. Een lang gesprek in de bus na een afdelingsuitje leverde ons genoeg voer op voor onderzoek ideeën. Ik ben hiermee hard aan de slag gegaan en het heeft me gebracht waar ik vandaag de dag ben. Ik ben je hier eeuwig dankbaar voor en ben ontzettend blij dat je al die jaren, ondanks de afstand, als copromotor bent aangebleven. Zie hier het resultaat!

Bij de onderzoeksprojecten waren vele mensen betrokken waarvan ik er toch drie bij name wil bedanken. Jelle Haringsma, Sonja Buzink en Vivian Ekkelenkamp.

Beste Jelle, ondanks onze zeer verschillende karakters konden we toch gezamenlijk vele goede ideeën creëren. Ik had al snel in de gaten dat ik jou apart moest spreken en om de tafel moest gaan zitten om tot goede projecten te komen. Als me dat gelukt was dan bleek jij vol inspirerende ideeën te zitten. Dit leidde ertoe dat we een aantal fantastische onderzoekslijnen hebben neergezet. Je bent een persoon zonder enige tijdsdruk, in mijn ogen vaak ad hoc en chaotisch. Ik denk dat we nog nooit op een afgesproken tijdstip begonnen zijn, doorgaans kwam je wat later aanzetten. Dat kon me vaak enorm frustreren maar had ook wel weer een andere kant. Ondanks een reeds uitlopend spreekuur kon je toch rustig een half uur met mij de zaken doorspreken als ik dat nodig had. Dat vond ik dan toch wel weer fijn.

Je bent een zeer begenadigd scopist en ik heb met veel plezier vele uren naast je gestaan om te kijken en te leren hoe jij de vaak lastige endoscopische problemen aanpakte. Veel daarvan pas ik nog altijd toe in mijn praktijk. Ik dank je voor een paar hele mooie jaren.

Lieve Sonja, de eerste twee onderzoeken deden wij samen. Zeer verschillende achtergronden; jij vanuit het industrial design vanuit de TU Delft en ik als jonge medicus. We konden veel van elkaar leren maar als het over onderzoek doen ging voelde ik me toch wel het groentje. Ik heb genoten van onze samenwerking die twee prachtige publicaties opleverden die in simulator-land veel geciteerd worden.

Lieve Vivian, mijn stok achter de deur! Het ziet er naar uit dat het me toch gelukt is vóór jou te promoveren. Wat heb ik jou zien groeien als onderzoeker! Helemaal onervaren kwam je uit Groningen om onderzoek te doen en uiteindelijk de opleiding tot MDL-arts binnen te slepen. Inmiddels ben je uitgegroeid tot een zeer zelfstandig onderzoeker. Een zeer moeilijke statistiek cursus gedaan die je ook geen windeieren heeft gelegd. Ik heb genoten van onze samenwerking en ben ook trots op je behaalde resultaten. Ik hoop stiekem dat ik straks als copromotor bij je promotie aan tafel mag zitten. Succes met de laatste loodjes en de aanstaande opleiding!

Ik wil de leden van de kleine commissie voor de beoordeling van mijn proefschrift hartelijk bedanken: prof. dr. M.J. Bruno, prof. dr. P.D Siersema en prof. dr. J.J.B van Lanschot.

Beste Marco, inmiddels heb je de scepter van Ernst overgenomen en dat doe je met veel verve. Ik meen een club MDL-artsen te zien waar iedereen zich happy voelt, zich als een teamspeler wil inzetten en iedereen de ruimte krijgt zich binnen zijn of haar vakgebied ten volle te ontplooien. Je hebt me enorm de ruimte gegeven om ten eerste mijn promotie af te ronden en me tevens te ontwikkelen in de geavanceerde endoscopie. Onze toekomstplannen zijn groots; de eerste stap in de vorm van een zeer succesvolle internationale endoscopie cursus op het gebied van geavanceerde endoscopie is daar een voorbeeld van! Ik verheug me op onze verdere gezamelijke toekomst.

Beste Peter, voordat je ook maar een woord gelezen had van het door jou te beoordelen proefschrift, probeerde je me al onder druk te zetten met een paar kritische opmerkingen. Dit soort opmerkingen kan ik erg waarderen en maken mij doorgaans scherper! Als editor-inchief van een van de voor mij belangrijkste vakbladen weet ik dat jij de lat hoog legt. Ik kijk dan ook uit naar een scherpe gedachtenwisseling.

Beste Jan, mijn proefschrift gaat over de eerste stappen in de endoscopie. Binnen mijn huidige werkzaamheden in de geavanceerde endoscopische resectietechnieken hebben wij veel meer raakvlakken en sparren wij bijna wekelijks over complexe pathologie. Je inzichten ervaar ik altijd als zeer verhelderend. Ik ben vereerd dat je in mijn commissie hebt plaatsgenomen en hoop dat we nog vele jaren vruchtbaar mogen samenwerken.

Ook de overige leden van de promotiecommissie wil ik van harte bedanken: prof. dr. J. van Saase, prof. dr. R.J. Stolker en prof. dr. M. Dinis-Ribeiro. Beste Jan en Robert Jan, jullie vertegenwoordigen een van de grootste opleidingsafdelingen in Nederland en een vakgebied waarin simulatortraining al op een veel hoger niveau staat. Dank voor jullie deelname aan wat de discussie ongetwijfeld naar een hoger niveau gaat tillen.

Dear Mário, it is my honour to have you as a friend, a colleague and today my highly learned opponent! The chairman of the educational committee of the ESGE in my PhD defense is for me the icing on the cake. We share a lot if ideas and working together so far has been a real pleasure for me. We have great plans for the future, also on the topics of advanced endoscopy training and education. The story continues... I am very much looking forward to it. But when are we finally going to have these caipirinha's?

En dan mijn paranimfen: Leon Moons en Wilco Lesterhuis! Jongens, wat vind ik het prachtig dat jullie mij terzijde willen staan. Leon, ik heb het altijd jammer gevonden dat je niet in Rotterdam gebleven bent en aangezien je het zeer naar je zin hebt in Utrecht zal dat er ook wel niet meer van komen. Ik genoot ook van onze fietstripjes in de vroege ochtend richting het ziekenhuis. Helaas heeft het ons niet de Adonis figuren opgeleverd waar we op gehoopt hadden... We hebben wat geintjes uitgehaald: posters ophangen van ons afdelingshoofd van de parkeergarage tot in het hele ziekenhuis op zijn 50ste verjaardag en aangeboden zijn screeningscolonoscopie te doen en solliciteren naar een praatje over Helicobacter pylori in het Atlantis resort in de Bahama's. Helaas gooide een vulkaanuitbarsting roet in het eten, anders waren we zeker gegaan, de koffers waren al gepakt zo getuige de foto's...

Wilco, jij lééft voor opleiden! Ik heb nog nooit een pre-pensionado met zoveel enthousiasme door de gangen van het ziekenhuis zien wervelen gevolgd door een horde van jonge honden als jij. Je energie is aanstekelijk en je bent altijd vrolijk. Ik vind het jammer dat je alweer van het toneel verdwenen bent, je hebt onze harten gestolen. Toen ik je vroeg mijn paranimf te zijn was je typerend als een kind zo blij en verrast. Al snel volgde je met de vraag: "Maar mag mijn vrouw dan ook komen? Zij wil me natuurlijk graag in dat apenpakkie zien en foto's maken..."

Ik wil graag al mijn collegae stafleden, AIOS, verpleegkundigen en secretaresses en iedereen van de MDL afdeling bedanken. Vaak genoeg kwam ik aanzetten met nog een extra scopietje voor een studie of weer een colo-tje afgepakt van een AIOS in het belang van de wetenschap. Hoewel het altijd druk is, kon het er toch altijd weer bij! Ik heb er heel wat worteltaarten voor gebakken en hoop dat het genoeg was, hoewel ik het antwoord daarop ook wel weer weet...

Dan is er nog een persoon die ik graag wil bedanken voor al haar geduld. Lieve, lieve Anna, kochanie, ook zonder jou was dit proefschrift er niet gekomen. Een zware last van onze beider schouders. Bedankt voor je liefde, zorgzaamheid en geduld. Op naar nog vele jaren szczęście, wolność i radość met z'n tweetjes en ons fantastische mannetje Stefan! Buziaki…

LIST OF CO-AUTHORS

Alphabetical order

S.M.B.I. Botden, MD PhD
Department of Surgery, Catharina Hospital Eindhoven, the Netherlands

S.N. Buzink, PhD
Faculty of Industrial Design Engineering, Delft University of Technology, the Netherlands

E. Dekker, MD PhD
Department of Gastroenterology and Hepatology, Academic Medical Center, Amsterdam, the Netherlands

F. van Delft, MD
Department of Gastroenterology and Hepatology, Academic Medical Center, Amsterdam, the Netherlands

V.E. Ekkelenkamp, MD
Department of Gastroenterology and Hepatology, Erasmus MC – University Medical Center Rotterdam, the Netherlands

P. Familiari, MD PhD
Digestive Endoscopy Unit, Gemelli University Hospital, Universita` Cattolica del Sacro Cuore, Rome, Italy

R.H.M. Goossens, PhD
Faculty of Industrial Design Engineering, Delft University of Technology, the Netherlands

J. Haringsma, MD
Department of Gastroenterology and Hepatology, Erasmus MC – University Medical Center Rotterdam, the Netherlands

A. Haycock, MD
Wolfson Unit for Endoscopy, St. Mark's Hospital, Imperial College London, UK

J. Heemskerk, MD
Department of Surgery, Catharina Hospital Eindhoven, the Netherlands

J.J. Jakimowicz, MD PhD
Department of Surgery, Catharina Hospital Eindhoven, the Netherlands

E.J. Kuipers, MD PhD
Department of Gastroenterology and Hepatology, Erasmus MC – University Medical Center Rotterdam, the Netherlands

R.A. de Man, MD PhD
Department of Gastroenterology and Hepatology, Erasmus MC – University Medical Center Rotterdam, the Netherlands

L. Petruzziello, MD
Digestive Endoscopy Unit, Gemelli University Hospital, Universita` Cattolica del Sacro Cuore, Rome, Italy

H. de Ridder, PhD
Faculty of Industrial Design Engineering, Delft University of Technology, the Netherlands

E.J. Schoon, MD PhD
Department of Gastroenterology and Hepatology, Catharina Hospital Eindhoven, the Netherlands

S. Thomas-Gibson, MD PhD
Wolfson Unit for Endoscopy, St. Mark's Hospital, Imperial College London, UK

R. Veenendaal, MD PhD
Department of Gastroenterology and Hepatology, Leiden University Medical Center, the Netherlands

LIST OF PUBLICATIONS

1. Komeda Y, Bruno MJ, **Koch AD**. EMR is not inferior to ESD for early Barrett's and EGJ neoplasia: an extensive review on outcome, recurrence and complication rates. Endoscopy International Open. Accepted for publication 2014.

2. Ekkelenkamp VE, **Koch AD**, Haringsma J, Kuipers EJ, de Man RA. Endoscopist-related factors contributing to high-quality colonoscopy: results of a Delphi survey. Perspect Med Educ. 2014 Jan;3(1):31-40.

3. Ekkelenkamp VE, **Koch AD**, Haringsma J, Poley JW, van Buuren HR, Kuipers EJ, de Man RA. Quality evaluation through self-assessment: a novel method to gain insight into ERCP performance. Frontline Gastroenterol. 2014 Jan;5(1):10-16.

4. **Koch AD**, de Man RA, Kuipers EJ. Response to Shaikh et al. Am J Gastroenterol. 2013 Jun;108(6):1012-3.

5. Sikkens EC, Cahen DL**, Koch AD**, Braat H, Poley JW, Kuipers EJ, Bruno MJ. The prevalence of fat-soluble vitamin deficiencies and a decreased bone mass in patients with chronic pancreatitis. Pancreatology. 2013 May-Jun;13(3):238-42.

6. Kerkhof M, Dewint P, **Koch AD**, van der Woude CJ. Endoscopic needle-knife treatment of refractory ileo-ascending anastomotic stricture. Endoscopy. 2013;45 Suppl 2 UCTN:E57-8.

7. Didden P, **Koch AD**. Endoscopic Assessment of Early Neoplasia in the Gastrointestinal Tract. European Medical Journal - Gastroenterology, 2012:1:45-52.

8. **Koch AD**, de Man RA, Kuipers EJ. Response to Singh. Am J Gastroenterol 2012; 107:1925-1926.

9. **Koch AD**, J Haringsma, EJ Schoon, RA de Man, EJ Kuipers. Competence Measurement During Colonoscopy Training: The Use of Self-Assessment of Performance Measures Am J Gastroenterol 2012;107:971–975.

10. **Koch AD**, Poley JW, Kuipers EJ, Bruno MJ. Advanced Therapeutic Endoscopy – The Continuing Shift towards Minimally Invasive Endoscopic Surgery. European Gastroenterology & Hepatology Review, 2011;7(3):189–93.

11. Haycock A, **Koch AD**, Familiari P, van Delft F, Dekker E, Petruzziello L, et al. Training and transfer of colonoscopy skills: a multinational, randomized, blinded, controlled trial of simulator versus bedside training. Gastrointest Endosc. 2010 Feb;71(2):298-307.

12. **Koch AD**, Haringsma J, Schoon EJ, de Man RA, Kuipers EJ. A second-generation virtual reality simulator for colonoscopy: validation and initial experience. Endoscopy. 2008 Sep;40(9):735-8.

13. **Koch AD**, van den Bosch HC, Bravenboer B. Epstein-Barr virus-associated cholecystitis. Ann Intern Med. 2007 Jun 5;146(11):826-7.

14. **Koch AD**, Buzink SN, Heemskerk J, Botden SM, Veenendaal R, Jakimowicz JJ, et al. Expert and construct validity of the Simbionix GI Mentor II endoscopy simulator for colonoscopy. Surg Endosc. 2008 Jan;22(1):158-62.

15. Buzink SN, **Koch AD**, Heemskerk J, Botden SM, Goossens RH, de Ridder H, et al. Acquiring basic endoscopy skills by training on the GI Mentor II. Surg Endosc. 2007 Nov;21(11):1996-2003.

16. **Koch AD**, Schoon EJ. Extensive jejunal diverticulosis in a family, a matter of inheritance? Neth J Med. 2007 Apr;65(4):154-5.

17. **Koch AD**, Holman ER, Versteegh MI, Klautz RJ, Somer ST, Bax JJ, et al. Reconstruction of the left ventricle after previous aneurysmectomy. Ann Thorac Surg. 2006 Apr;81(4):1495-7.

18. **Koch AD**, Nicolai JP, de Vries J. Breast cancer and the role of breast size as a contributory factor. Breast. 2004 Aug;13(4):272-5.

19. Schoof PH, **Koch AD**, Hazekamp MG, Waterbolk TW, Ebels T, Dion RA. Bovine jugular vein thrombosis in the Fontan circulation. J Thorac Cardiovasc Surg. 2002 Nov;124(5):1038-40.

CURRICULUM VITAE

Arjun Dave Koch werd op 28 februari 1973 geboren te Harlingen. In 1991 behaalde hij het VWO eindexamen aan de Rijksscholengemeenschap Magister Alvinus te Sneek. Vervolgens studeerde hij geneeskunde aan de Rijksuniversiteit Groningen. Hier behaalde hij in 1997 zijn doctoraal examen en in 1999 zijn artsenbul. Een keuze coschap werd verricht op de afdeling Plastische en Reconstructieve chirurgie van het Academisch Ziekenhuis Groningen onder leiding van Prof. dr. P.H. Robinson. In samenwerking met Prof. dr. J.P.A. Nicolai, plastisch chirurg, en Dr. J. de Vries namens de afdeling Chirurgische Oncologie werd onderzoek verricht naar de relatie tussen borstkanker en borstgrootte wat leidde tot een publicatie in Breast. Na het behalen van zijn artsenbul werkte hij een aantal jaren als AGNIO op de afdelingen Cardio-thoracale chirurgie van het Academisch Ziekenhuis Groningen en Leids Universitair Medisch Centrum. Hierna koos hij voor de opleiding interne geneeskunde en belandde in het Catharina Ziekenhuis in Eindhoven bij opleider Dr. B. Bravenboer, internist-endocrinoloog. Hier kwam hij voor het eerst in zijn carrière in aanraking met het vak maagdarmlever ziekten en was het voornemen over te stappen naar de MDL-opleiding snel genomen. Onder begeleiding van Dr. E.J. Schoon werd een onderzoekslijn op endoscopie simulatoren opgezet. Dit leidde ertoe dat hij in 2006 de overstap maakte naar de opleiding tot MDL-arts in het Erasmus MC te Rotterdam. De opleiding werd in 2010 afgerond en sindsdien is hij als MDL-arts en staflid aan de afdeling MDL-ziekten van het Erasmus MC verbonden met het aandachtsgebied geavanceerde endoscopie. Het promotie onderzoek in de lijn van competentie gericht trainen, waarvan de eerste stappen in Eindhoven waren gezet, werd doorgezet onder begeleiding van Prof. dr. E.J. Kuipers als promotor en Dr. E.J. Schoon en Dr. R.A. de Man als copromotoren.

Sinds 2011 is hij tevens manager van de ESE, de Erasmus School of Endoscopy, en in die hoedanigheid organiseert hij diverse trainingen voor zowel de beginnende endoscopist als ervaren scopisten op het gebied van geavanceerde endoscopische resecties. Hij is daarnaast actief in de commissie voor de ontwikkeling van het Nederlandse e-portfolio voor MDL-artsen in opleiding en sinds 2013 lid van de ESGE Educational Committee. Tevens is hij als Toetsing Coördinator MDL betrokken bij de ontwikkeling en uitvoering van de toetsing van endoscopisten voor de toetreding als screenings-endoscopisten voor het nationale bevolkingsonderzoek naar dikkedarmkanker. Hij is gehuwd met Anna in 2010 en in 2011 is hun zoon Stefan geboren.

PHD PORTFOLIO

presentations

2006	NVGE Veldhoven: The GI Mentor II simulator: validation and learning curves
	UEG Berlin: Expert and Construct Validity of the GI Mentor II Simulator for Colonoscopy (poster presentation)
	UEG Berlin: Acquiring basic endoscopy skills by training on the GI Mentor II (poster presentation)
2007	NVGE Veldhoven: Teaching Endoscopy - Validation of a second generation Simulator
	UEGW Paris: Validation of a Novel Virtual Reality Simulator for Colonoscopy (poster presentation)
	Waddensymposium Terschelling: Chronic mesenteric ischemia
2008	National GI Trainers meeting, Netherlands: Skills Training
2009	UEG London: Patient-based assessment of prolonged colonoscopy simulator training
	NVGE Veldhoven: Patient-based assessment of prolonged colonoscopy simulator training
	Regional GI meeting Rotterdam: Teaching Endoscopy - The SAFE study
2010	DDW New Orleans, ASGE presidential plenary session: Self-assessment in Colonoscopy - a novel tool for assessment of skills
	DDW New Orleans: Self-assessment for ERCP - a promising tool for trainees (poster presentation)
	World Helicobacter conference, Rotterdam, the Netherlands: Endoscopic management of early gastric cancer
	National GI Trainers meeting Utrecht: Teaching Endoscopy – From Novice to Competence and Excellence
	Regional GI meeting UMCG Groningen: Stenting benign and malignant biliary strictures
2011	Regional GI meeting Rotterdam: Endoscopic management of early GI neoplasia
2012	ESO, Almaty, Kazakhstan: Evidence for optimal screening strategy
	Advanced endoscopy workshops, Rotterdam: EMR in Barrett’s

2013 UEG Learning Area Berlin: Endoscopic treatment of peptic ulcer bleeds

EAGEN/ESGE Advanced endoscopy workshops, Rotterdam: EMR in Barrett's

EAGEN/ESGE Advanced endoscopy workshops, Rotterdam: ESD in the esophagus

ESDE Conference, Rotterdam: EMR in Barrett's

2014 EAGEN/ESGE Advanced endoscopy workshops, Braga, Portugal: ESD in the esophagus

NVGE Veldhoven: Meten is weten - ERCP kwaliteitsregistratie in Nederland

book chapters

J Rothbarth, AD Koch, JJB van Lanschot. Endoscopic and surgical treatment of early gastric cancer. In: Griffin SM, Raimes SA, editors. Oesophagogastric Surgery: A Companion to Specialist Surgical Practice. 5th ed. Saunders Elsevier; 2014

memberships

Dutch Society of Gastroenterology (NVGE)
Dutch Society of Gastroenterologists (NVMDL)
Dutch society of Hepatology (NVH)
Fellow of the European Board of Gastroenterology and Hepatology (EBGH)
European Society for Gastrointestinal Endoscopy (ESGE)

positions

Staff member department of Gastroenterology and Hepatology at the Erasmus MC – University Medical Center Rotterdam (2010 -)
Manager of the ESE – Erasmus School of Endoscopy (2010 -)
ESGE – Educational committee (2013 -)
Committee for the development of the Dutch e-portfolio for gastroenterologists in training (2010 -)
Examiner for the Dutch Colorectal Cancer Screening Program (TCMDL) (2012-2014)
Committee for the Dutch prospective quality registry on ERCP (2013 -)

ISBN 978-94-6169-519-2