Clinical Gastroenterology and Hepatology 2014;12:1611–1623

SYSTEMATIC REVIEWS AND META-ANALYSES

Fasiha Kanwal, Section Editor

Effects of Simulation-Based Training in Gastrointestinal Endoscopy: A Systematic Review and Meta-analysis

Siddharth Singh,* Robert E. Sedlack,* and David A. Cook‡,§

*Division of Gastroenterology and Hepatology, ‡Division of General Internal Medicine, Mayo Clinic, Rochester; and §Office of Education Research, Mayo Medical School, Rochester, Minnesota

BACKGROUND & AIMS: **Simulation-based training (SBT) in gastrointestinal endoscopy has been increasingly adopted by gastroenterology fellowship programs. However, the effectiveness of SBT in enhancing trainee skills remains unclear. We performed a systematic review with a meta-analysis of published literature on SBT in gastrointestinal endoscopy.**

METHODS: **We performed a systematic search of multiple electronic databases for all original studies that evaluated SBT in gastrointestinal endoscopy in comparison with no intervention or alternative instructional approaches. Outcomes included skills (in a test setting), behaviors (in clinical practice), and effects on patients. We pooled effect size (ES) using random-effects meta-analysis.**

RESULTS: **From 10,903 articles, we identified 39 articles, including 21 randomized trials of SBT, enrolling 1181 participants. Compared with no intervention (n = 32 studies), SBT significantly improved endoscopic process skills in a test setting (ES, 0.79; n = 22), process behaviors in clinical practice (ES, 0.49; n = 8), time to procedure completion in both a test setting (ES, 0.79; n = 16) and clinical practice (ES, 0.75; n = 5), and patient outcomes (procedural completion and risk of major complications; ES, 0.45; n = 10). Only 5 studies evaluated the comparative effectiveness of different SBT approaches; which provided inconclusive evidence regarding feedback and simulation modalities.**

CONCLUSIONS: **Simulation-based education in gastrointestinal endoscopy is associated with improved performance in a test setting and in clinical practice, and improved patient outcomes compared with no intervention. Comparative effectiveness studies of different simulation modalities are limited.**

Keywords: Simulation; Gastrointestinal Endoscopy; Outcomes; Education.

Traditional training in gastrointestinal endoscopy is based on the apprenticeship model (ie, trainees learn basic endoscopic skills under the supervision of experienced endoscopists in clinical practice). However, in light of ethical and medicolegal concerns for patient comfort and safety, as well as the negative short-term financial impact of teaching endoscopy to trainees, there is an increasing shift to simulation-based training (SBT) in gastrointestinal endoscopy.[1] SBT is an attractive alternative for teaching psychomotor and perceptual skills, offering an environment that avoids time pressures and patient safety risks and enables systematic variation of the clinical scenario. In fact, current guidelines from the American Council for Graduate Medical Education mandate the incorporation of SBT in all gastroenterology fellowship programs.[2]

Previous reviews have offered some insights into the effectiveness of SBT in gastrointestinal endoscopy, but these reviews have been limited by the lack of a systematic search, incomplete assessment of study quality, and an absence of quantitative pooling to derive best estimates of effect of these interventions on the trainees' endoscopic skills.[3,4] A recent review from the Cochrane Collaboration focused only on randomized controlled trials of computed-based endoscopy training, and included only 13 trials with 278 participants.[5] Nonrandomized studies, single-arm pre- vs postintervention comparisons, and on ex vivo animal models and mechanical simulation models were not included, and hence the review did not synthesize the available evidence comprehensively. A comprehensive review and

Abbreviations used in this paper: CI, confidence interval; EGD, esophagogastroduodenoscopy; ERCP, endoscopic retrograde cholangiopancreatography; ES, effect size; SBT, simulation-based training.

1542-3565/$36.00
http://dx.doi.org/10.1016/j.cgh.2014.01.037

synthesis would allow an objective assessment of the effectiveness of SBT, as compared with no intervention, in improving procedural skills and effects on patients, enable identification of appropriate instructional design features, and identify areas in simulation-based education that require further investigation.

Hence, we sought to identify and summarize, both quantitatively and qualitatively, all comparative studies of SBT in gastrointestinal endoscopy (diagnostic and therapeutic esophagogastroduodenoscopy [EGD], colonoscopy, flexible sigmoidoscopy, endoscopic retrograde cholangiopancreatography [ERCP], and endoscopic ultrasound) through a systematic review and meta-analysis of published literature.

Methods

This study was a planned subanalysis of data collected as part of a comprehensive review of simulation-based education.[6] The study was planned, conducted, and reported in adherence to Preferred Reporting Items for Systematic Reviews and Meta-Analysis standards of quality for reporting meta-analyses.[7] Our general methods have been described in detail previously[6]; we summarize them briefly later.

Questions

We sought to answer the following questions: (1) what is the effectiveness of technology-enhanced simulation for training in gastrointestinal endoscopy on trainee knowledge, procedural skills in a simulation/training environment, performance skills with actual patients, and effects on patient outcomes, and (2) what instructional design features are associated with improved outcomes in trainee performance? We defined technology-enhanced simulation as an educational tool or device with which the learner physically interacts to mimic an aspect of clinical care.[6]

Study Eligibility

We included studies involving health professional learners at any stage in training or practice that investigated the use of technology-enhanced simulation to learn gastrointestinal endoscopy, in comparison with the following: (1) no intervention (ie, a control arm or pre-intervention assessment), (2) a nonsimulation training activity, or (3) an active alternative SBT modality. Both single-group pretest-posttest and 2-group randomized and nonrandomized studies, focusing on the educational outcomes of SBT in therapeutic EGD, colonoscopy, flexible sigmoidoscopy, ERCP, and endoscopic ultrasound, were included. We did not exclude studies based on outcome, year, or language of publication. However, studies that focused only on simulation-based assessment (ie, a model's ability to assess procedural skills) were excluded.

Study Identification

We performed a systematic literature search of MEDLINE, EMBASE, CINAHL, PsycINFO, ERIC, Web of Science, and Scopus, from inception to May 11, 2011, with the help of an experienced librarian. Our full search strategy has been published previously.[6] We searched for omitted articles by reviewing the reference lists of all included articles, technical reviews from the American Society of Gastrointestinal Endoscopy,[8] the September 2006 edition of *Gastrointestinal Endoscopy Clinics of North America*, which was dedicated to Endoscopy Simulators for Training and Assessing Skills (which included the consensus statement of the First International Conference on Endoscopy Simulation),[4,9–14] and several published reviews of health professions simulation.[3,5,15] Finally, we searched the full table of contents of 2 journals devoted to health professions simulation (*Simulation in Healthcare* and *Clinical Simulation in Nursing*).

During the peer-review phase, we identified recently published articles by searching PubMed using the terms simulat* AND (egd OR endoscopy OR colonoscopy OR ercp) from a date range of January 11, 2011, to December 19, 2013. We retrieved 765 articles. A single author reviewed all of these studies, applying the inclusion criteria noted earlier and extracting key information from eligible studies.

Study Selection

Study selection was performed in 2 stages. In the first stage, we identified all studies of technology-enhanced simulation for health professional education using the search described earlier. Two reviewers independently screened all titles and abstracts to exclude studies that did not address the research question of interest. The full texts of the remaining articles were reviewed for definitive inclusion or exclusion, again independently and in duplicate. We resolved conflicts by consensus. Chance-adjusted interrater agreement for study inclusion at the first step was substantial (intraclass correlation coefficient, 0.69).[16] In the second stage, 2 investigators trained in gastroenterology reviewed the studies addressing technology-enhanced simulation to specifically identify studies focused on gastrointestinal endoscopy training; the κ-coefficient of agreement between the 2 investigators at this stage was 0.95.

Data Extraction

We abstracted information independently and in duplicate for all variables in which reviewer judgment was required, and resolved conflicts by consensus. Foreign-language articles were translated before data abstraction. By using a data abstraction form, we abstracted information on the training level of learners, clinical topic, method of group assignment, outcomes

reported, and methodologic quality. Methodologic quality was graded using the Medical Education Research Study Quality Instrument[17] and an adaptation of the Newcastle–Ottawa scale for cohort studies.[18] We also coded key instructional design features of the intervention.[19]

We abstracted information separately for learning outcomes of trainee satisfaction, medical knowledge, skills (performance in a test setting), behaviors (performance with real patients in clinical practice), and direct effects on patients (procedure completion rates, major complications, patient discomfort). Skills and behaviors were classified further as measures of the endoscopy process (eg, economy of movement, minor errors, or global ratings in a test setting or in clinical practice), and time (time to complete the procedure in a test setting or in clinical practice). Outcomes could be measured subjectively (learner or patient self-report) or objectively (observer rating, objective assessment of acquired clinical skills). We converted reported results to a standardized mean difference (Hedges' g effect size) using methods described previously.[6] For articles containing insufficient information to calculate an effect size (ES) we requested information from the investigators.

Data Synthesis

We grouped studies according to the comparison arm, namely, no intervention (comparison with another group receiving no endoscopy training, or the same group at baseline [single-arm pre/poststudy]), nonsimulation intervention (comparison with a group receiving training with usual methods [eg, real patients]), or alternate simulation intervention (comparison with another SBT approach). For quantitative synthesis, we pooled, using meta-analysis, the results of studies comparing simulation training with no intervention. We also planned meta-analysis of studies, making comparison with another active instructional intervention whenever 2 or more studies made a sufficiently similar comparison. We planned sensitivity analyses that excluded nonrandomized studies and studies with imprecise ES calculations (calculated using *P* value upper limits or imputed standard deviations). We also performed subgroup analyses based on the gastrointestinal endoscopic procedure (eg, EGD, colonoscopy), modality of SBT (virtual reality vs ex vivo models), training level of participants (medical students, postgraduate trainees including residents and fellows, practicing endoscopists), and study quality.

We quantified heterogeneity across studies using the inconsistency index (I^2 statistic),[20] which estimates what proportion of total variation across studies was not owing to chance. I^2 values greater than 50% indicated significant inconsistency. Anticipating significant heterogeneity in the analyses, we used random-effects models to pool weighted ES. We used SAS 9.3 (SAS Institute, Cary, NC) for all analyses. A *P* value of less than .05 was considered statistically significant; clinical significance of ES was interpreted in relation to Cohen's ES classifications (>0.8 was considered a large ES, 0.5–0.8 was considered moderate, and 0.2–0.5 was considered small).[21] We evaluated the possibility of publication bias, both qualitatively by visual inspection of funnel plots and quantitatively using the Egger asymmetry test, recognizing that these procedures can be misleading in the presence of significant heterogeneity.

We also conducted a nonquantitative synthesis of evidence focused on key instructional design features. We reviewed all comparative-effectiveness studies (studies comparing one simulation modality with another active simulation-based intervention) to inductively identify salient themes. We used, as an initial guide, the 10 key features of high-fidelity simulation identified by Issenberg et al,[22] but we also remained open to the possible presence of other design features.

Results

Trial Flow

By using our broad search strategy, we identified 10,903 articles, from which we identified 985 comparative studies of technology-enhanced simulation for health professional training (Figure 1). From this set, we identified 39 articles focused on gastrointestinal endoscopy training. Thirty-two of these studies compared SBT with no intervention, 2 studies compared SBT with nonsimulation training, and 5 studies compared one modality of SBT with another simulation-based instructional approach. One article was published in German.[23] Through an updated search during the peer-review phase, we identified 8 additional studies, including 6 studies making a comparison with no intervention, one study comparing simulation vs nonsimulation instruction, and one study comparing high vs low fidelity. Key findings from these studies are reported in the Supplementary Results and are summarized in Supplementary Table 1; these were not included in quantitative synthesis later.

Study Characteristics

Study characteristics are detailed in Table 1 and summarized in Table 2. Postgraduate trainees with no endoscopy exposure or early in their endoscopic training were the most frequent learners (n = 32 studies). Most of these learners were fellows in gastroenterology (n = 20 studies), or residents (n = 19 studies) in internal medicine (n = 16 studies) and/or general surgery (n = 8 studies). Six studies involved medical students. Colonoscopy and EGD were the most common endoscopic skills taught in SBT.

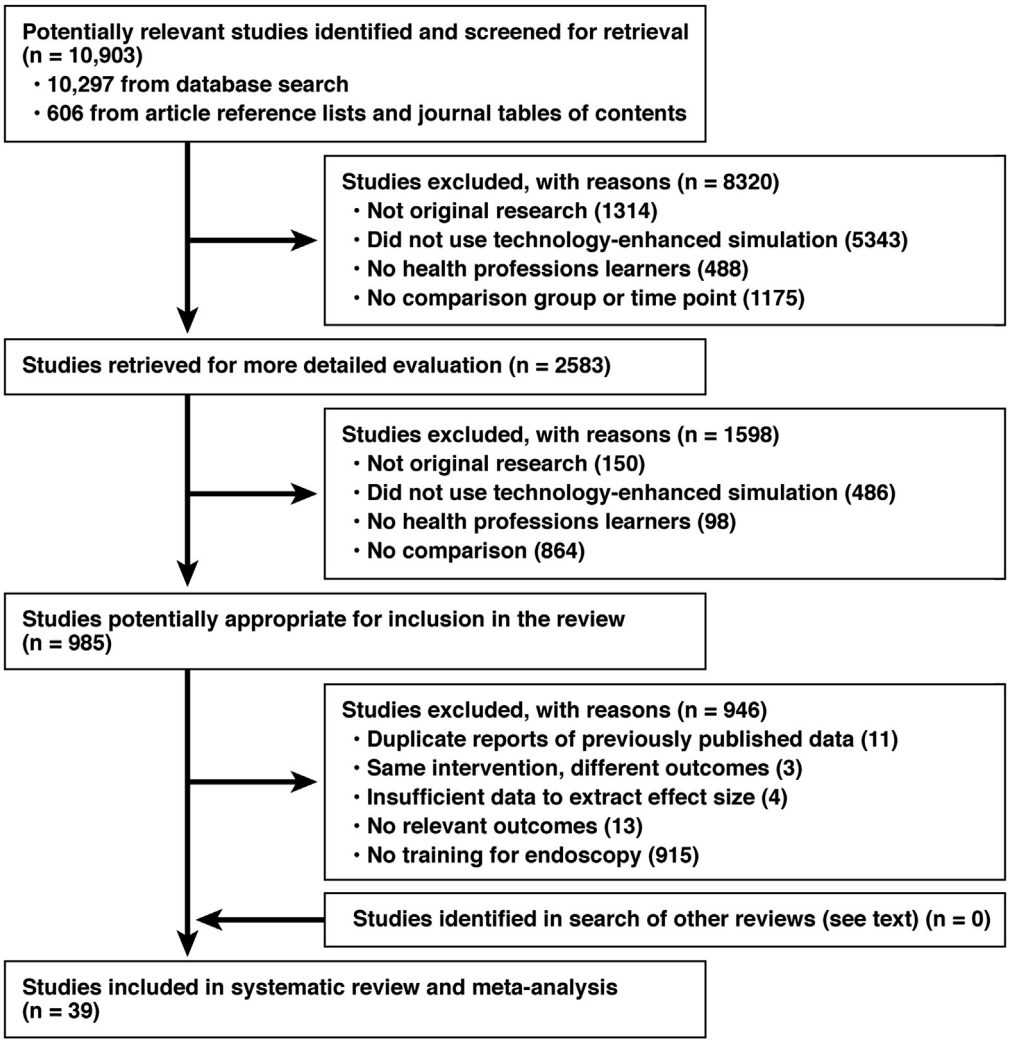

Figure 1. Trial flow.

Simulation modalities varied widely. Virtual-reality or computer-based SBT was the most common modality studied (n = 29 studies). Such simulators typically consist of a proxy flexible gastrointestinal endoscope introduced into an interface device that transmits movements to a computer. The computer displays images of the upper and lower intestinal tracts. Many systems provide some form of force feedback and reproduce events such as patient discomfort. Ten studies used ex vivo animal organ models to enhance tissue realism in SBT, most commonly involving porcine stomach. Two studies used part-task trainers (specifically designed to teach a specific skill) for SBT. Common instruction design features[19,22] included cognitive interactivity (ie, training that promotes learners' cognitive engagement using strategies such as multiple repetitions, feedback, task variation, or intentional task sequencing) in 32 studies, distributed practice (ie, training spread over a period of time, which for this review we counted as present for interventions that involved >1 day of simulation training) in 24 studies, clinical variation (ie, variation in the clinical context, eg, multiple different patient scenarios in 21 studies), feedback (ie, information on performance provided to the learner by the instructor, a peer, or a computer, either during or after the simulation activity) in 14 studies, and mastery learning (ie, a training model in which learners must attain a clearly defined standard of performance before qualifying or advancing to the next task) in 7 studies. Curricular integration (ie, incorporation of the simulation intervention as an integral part of the curriculum or training program) was performed in only 2 studies.

The outcomes studied included computer-generated estimates of performance, subjective or objective assessments by observers, and satisfaction questionnaires. Fourteen studies assessed performance in the context of gastrointestinal endoscopy of real patients. Most of the remaining studies used simulation-based outcome measures including combinations of dexterity, accuracy, and speed (economy of performance). Two studies assessed reaction/satisfaction of trainees after simulation training, and one study assessed knowledge.

Study Quality

Study quality is summarized in Tables 1 and 3. Thirteen studies used a single-group pretest-posttest design. Of the 26 two-group studies, 21 used randomized

Table 1. Description of Studies Included in This Systematic Review of SBT in Gastrointestinal Endoscopy

Study	Trainee, N, level	Study design	Comparison	Task	Modality	Instruction features	Outcomes reported	Follow-up[a]	Outcome blind[b]	Outcome objective	MERSQI total	NOS total
Tuggy, 1998[41]	10; PG	RCT	NI	FS	VR		ST, SP		-	Obj	11.5	3
Ferlitsch et al, 2002[42]	13; PG	RCT	NI	EGD, colon	VR	CI, CV, DP, Mast	ST, SP		Blind	Obj	12.5	3
Gerson and Van Dam, 2003[27]	16; PG	2NR	OE	FS	VR	CI, CV, DP, MLS, Reps	BT, BP, P	High	-	Obj	13.5	2
Kneebone et al, 2003[43]	7; RN	1PP	NI	FS	VR	CI	ST, SP		Blind	Obj	10	1
Neumann et al, 2003[44]	56; PG	1PP	NI	EGD	Organ	CI, DP, FB	SP	High	-	Obj	14	1
Neumann et al, 2003[45]	25; MS, PG, MD	1PP	NI	EGD	Organ	CI, DP, FB, Mast	SP		-	Obj	12	0
Di Giulio et al, 2004[46]	22; PG	RCT	NI	EGD	VR	CI, CV, DP	BT, P		-	Obj	14	3
Eversbusch and Grantcharov, 2004[47]	20; O	RCT	NI	Colon	VR	CI	ST, SP		Blind	Obj	12.5	4
Hochberger et al, 2004[31]	207; PG, MD, RN, O	2NR	SS	EGD	Organ	CI, FB	R	High	-	Subj	10	2
Mahmood and Darzi, 2004[25]	26; PG, MD	1PP	NI	Colon	VR		ST, SP	High	Blind	Obj	11	2
Sedlack et al, 2004[48]	38; PG, MD	RCT	NI	FS	VR	CI, Curr, CV	P	High	Blind	Subj	13	4
Sedlack and Kolars, 2004[49]	8; PG	RCT	NI	Colon	VR	CI, CV, DP, Reps	BT, BP, P		-	Obj	13	2
Ahlberg et al, 2005[50]	12; PG	RCT	NI	Colon	VR	CI, CV, DP, FB, Mast, MLS	BT, P		Blind	Obj	15	4
Hochberger et al, 2005[51]	28; PG	RCT	NI	EGD	Organ	CI, CV, DP, FB	ST, SP, P	High	Blind	Obj	16	5
Kiesslich et al, 2005[23]	100; MD	2NR	SS	EGD	VR, Organ	CI, FB	R	High	Blind	Obj	11	3
Matthes et al, 2005[52]	8; PG	1PP	NI	EGD	Organ	CI, FB	ST, SP	High	-	Obj	12	1
Neumann et al, 2005[53]	58; MS, PG	1PP	NI	EGD	Organ	CI, DP, FB	SP	High	Blind	Obj	14	2
Cohen et al, 2006[54]	49; PG	RCT	NI	Colon	VR	CI, CV, DP, Reps	BP, P	High	Blind	Obj	15	6
Maiss et al, 2006[55]	35; PG	RCT	NI	EGD	Organ	CI, CV, DP, FB	ST, SP		Blind	Obj	13	5
Ritter et al, 2006[56]	16; MS	1PP	NI	FS	VR	CI	ST, SP		Blind	Obj	10.5	2
Thomson et al, 2006[57]	14; PG	2NR	NI	Colon	VR	CI, Curr, CV, DP	BP, P	High	-	Obj	13	4
Buzink et al, 2007[58]	30; PG	1PP	NI	Colon	VR	CI, CV, DP, Reps	ST, SP	High	Blind	Obj	11	2
Maiss et al, 2007[59]	27; MS, PG	RCT	NI	EGD	Organ	CI, DP, FB	ST, SP	High	Blind	Obj	12.5	5
Park et al, 2007[60]	28; PG	RCT	NI	Colon	VR	CI, CV	BP, P	High	Blind	Obj	16	4
Sedlack, 2007[24]	8; PG	RCT	NI	EGD	VR	CV, DP, Reps	BP	High	-	Obj	16	4
Thomas-Gibson et al, 2007[26]	21; PG	1PP	NI	Colon	VR	CI, CV, DP, MLS	K, ST, SP, BP	High	Blind	Obj	12.5	2
Yi et al, 2007[61]	9; PG	1PP	NI	Colon	VR	CV, Mast, Reps	ST, SP		Blind	Obj	10	1
Shirai et al, 2008[62]	20; PG	RCT	NI	EGD	VR	CV, DP	SP		Blind	Obj	12.5	4
Yi et al, 2008[63]	11; PG	2NR	NI	Colon	VR	CI, CV, DP, Mast, Reps	BT, BP, P	High	-	Subj	10	2
Haycock et al, 2009[64]	28; PG, MD	RCT	NI	EGD	Organ	CI, CV	SP	High	Blind	Obj	14.5	5
Walsh et al, 2009[30]	30; MS	RCT	SS	Colon	Box	CI, FB, Reps	ST, SP	High	Blind	Obj	13.5	4

Table 1. Continued

Study	Trainee, N, level	Study design	Comparison	Task	Modality	Instruction features	Outcomes reported	Follow-up[a]	Outcome blind[b]	Outcome objective	MERSQI total	NOS total
Yi et al, 2009[65]	9; PG	1PP	NI	Colon	VR	CI, CV, Mast, Reps	ST, SP		Blind	Obj	10	1
Buzink et al, 2010[32]	29; MS	RCT	SS	Colon	VR	CV, DP	ST		Blind	Obj	11.5	3
Ferlitsch et al, 2010[66]	28; PG	RCT	NI	EGD	VR	CI, CV, DP, MLS, Reps	BT, BP, P	High	-	Obj	14	4
Haycock et al, 2010[28]	40; PG, RN, O	RCT	OE	Colon	VR	CI, CV, DP, FB	ST, SP, BT, BP, P	High	Blind	Obj	15	5
Haycock et al, 2010[67]	28; PG	RCT	NI	EGD	Box		SP	High	Blind	Obj	12.5	4
Kruglikova et al, 2010[68]	30; PG, RN	1PP	NI	Colon	VR	CI, DP, FB	ST, SP		Blind	Obj	11	1
Kruglikova et al, 2010[29]	22; PG	RCT	SS	Colon	VR	CI, DP, FB, Reps	ST, SP	High	Blind	Obj	12.5	5
Luursema et al, 2010[69]	15; O	1PP	NI	Colon	VR	CI, DP	ST, SP	High	Blind	Obj	11	2

BP, process behavior; Box, box-trainer; BT, time behavior; CI, cognitive interactivity present; colon, colonoscopy; curr, curricular integration; CV, clinical variation present; DP, distributed practice (>1 day); FB, feedback high; FS, flexible sigmoidoscopy; K, knowledge; mast, mastery learning; MD, practicing physician; MERSQI, Medical Education Research Study Quality Instrument (maximum score 18); MLS, multiple learning strategies; MS, medical student; NI, no intervention; NOS, modified Newcastle–Ottawa scale (maximum score 6); 2NR, nonrandomized 2-group study; O, other health professional; obj, objective assessment; OE, other (nonsimulation) education; organ, ex vivo animal organ; P, patient effects; PG, postgraduate physician trainee (resident); 1PP, 1 group pre-post study; R, reaction (satisfaction); RCT, randomized 2-group study; reps, many repetitions; RN, registered nurse; SP, process skill; SS, simulation-simulation comparison; ST, time skill; subj, self-report of highest-level outcome; VR, virtual reality.

[a]High represents ≥75% follow-up evaluation for any outcome.

[b]Blinded assessment of highest-level outcome.

Table 2. Summary of Key Features of Included Studies

Study characteristic	Level	Studies, n (no. of participants[a])
All studies		39 (1181)
Study design	2-group	26 (871)
	1-group (pretest-posttest)	13 (310)
Group allocation	Randomized	21 (523)
Comparison	No intervention	32 (737)
	Nonsimulation training	2 (56)
	Alternate simulation training	5 (388)
Participants[b]	Medical students	6 (111)
	Physicians postgraduate training	32 (632)
	Physicians in practice	6 (204)
	Nurse	4 (92)
	Other/ambiguous/mixed	5 (142)
Task[b]	EGD	16 (691)
	Colonoscopy	19 (416)
	Flexible sigmoidoscopy	4 (71)
	ERCP/EUS	0
Simulation modalities[b]	Virtual reality	28 (651)
	Part-task model/box trainer	2 (58)
	Live animal	1 (36)
	Animal tissue (cadaveric)	10 (572)
Outcomes[b]	Satisfaction	2 (307)
	Knowledge	1 (21)
	Skill: time[c]	20 (399)
	Skill: process[c]	25 (585)
	Behavior: time[d]	7 (133)
	Behavior: process[d]	10 (210)
	Patient effects[e]	12 (267)
Quality	Newcastle–Ottawa ≥4 points	17 (455)
	MERSQI ≥12 points	26 (682)

EUS, endoscopic ultrasound; MERSQI, Medical Education Research Study Quality Instrument.

[a]Numbers reflect the number enrolled, except for the Outcomes section, which reflect the number of participants who provided observations for analysis.

[b]The number of studies and trainees in some subgroups may total to more than the number for all studies because several studies included more than 1 trainee group or simulation modality, fit within more than 1 task, or reported more than 1 outcome.

[c]Skills: subjective (eg, learner self-report) or objective (eg, faculty ratings, or objective tests of clinical skills such as computer-scored technique in a virtual reality simulator) assessments of learners' ability to demonstrate a procedure or technique in an educational setting (typically a simulation task). We further classified skills as measures of time (how long it takes a learner to complete the task) and process (eg, global rating scales, efficiency, or minor errors, successful completion of the task, evaluation of the finished product, or major errors that would impact a real patient's well-being).

[d]Behaviors: Subjective (eg, learner or patient self-report) or objective (eg, chart audit or faculty ratings) assessments of behaviors in practice or effects on patients (such as medical errors). We used a classification system similar to that used for skills, with time and process measures counted as behaviors (eg, procedure time, economy of movement, and so forth).

[e]Patient effects: complications, patient discomfort, or procedure completion rates.

group allocation. Twenty-seven studies used blinded outcome assessment. Fifteen studies lost more than 25% of enrolled participants before follow-up evaluation or failed to report the number of participants included in the analysis. Validity evidence to support outcome assessments was reported infrequently: 14 studies reported relations with other variables, 6 studies reported

Table 3. Summary of Quality of Included Studies

Scale item	Subscale (points if present)	No. (%) present (N = 39)
MERSQI[a]		
Study design (maximum 3)	1-group pre-post (1.5)	13 (33)
	Observational 2-group (2)	5 (13)
	Randomized 2-group (3)	21 (54)
Sampling: no. of institutions (maximum, 1.5)	1 (0.5)	28 (72)
	2 (1)	1 (3)
	>2 (1.5)	10 (25)
Sampling: follow-up (maximum 1.5)[b]	<50% or not reported (0.5)	13 (33)
	50%–74% (1)	2 (5)
	≥75% (1.5)	24 (62)
Type of data: outcome assessment (maximum, 3)	Subjective (1)	3 (8)
	Objective (3)	36 (92)
Validity evidence (maximum, 3)	Content (1)	6 (15)
	Internal structure (1)	4 (10)
	Relations to other variables (1)	13 (33)
Data analysis: appropriate (maximum, 1)	Appropriate (1)	37 (95)
Data analysis: sophistication (maximum, 2)	Beyond descriptive analysis (2)	39 (100)
Highest outcome type (maximum, 3)	Reaction (1)	2 (5)
	Knowledge, skills (1.5)	23 (59)
	Behaviors (2)	2 (5)
	Patient/health care outcomes (3)	12 (31)
Newcastle–Ottawa scale (modified)[c]		
Representativeness of sample	Present (1)	3 (8)
Comparison group from same community	Present (1)	26 (67)
Comparability of comparison cohort, criterion A[d]	Present (1)	21 (54)
Comparability of comparison cohort, criterion B[d]	Present (1)	14 (36)
Blinded outcome assessment	Present (1)	27 (69)
Follow-up high[e]	Present (1)	26 (67)

[a]The mean (SD) Medical Education Research Study Quality Instrument (MERSQI) score was 12.6 (1.8), and the median score (range) was 12.5 (10–16).
[b]Proportion of participants followed till end of study.
[c]The mean (SD) Newcastle–Ottawa Scale score was 3.0 (1.5), and the median score (range) was 3 (0–6).
[d]Comparability of cohorts: criterion A was present if the study (1) was randomized, or (2) controlled for a baseline learning outcome; criterion B was present if (1) a randomized study concealed allocation, or (2) an observational study controlled for another baseline trainee characteristic.
[e]Follow-up high refers to >75% follow-up evaluation for any outcome.

content evidence, and only 4 studies reported score reliability.

Synthesis

Effectiveness in comparison with no training. Thirty-two studies compared SBT with no intervention (either single-group comparison with baseline, or in comparison with a control group without formal training). For the 22 studies reporting process skills (ratings of technique, economy of movement, completion rate such as cecal intubation, and so forth, in a test setting), SBT resulted in significantly better performance with a moderate-large pooled ES of 0.79 (95% confidence interval [CI], 0.54–1.05; $P < .001$) (Figure 2). Likewise, time to procedural completion (time skill) was shorter in trainees who received SBT, with a pooled ES of 0.79 (95% CI, 0.49–1.05; $P < .001$; 16 studies). However, between-study inconsistency was high in both analyses (I^2, 70% for process skill and 61% for time skill) and individual ES in studies ranged from -0.32 to 3.23 for process skills and from -0.25 to 2.74 for time skills (one study in each analysis showed a negative ES). SBT also was associated with small-moderate effects on process behaviors (ES, 0.49; 95% CI, 0.21–0.77; $P < .001$; 8 studies) and time to procedure completion in clinical practice (time behaviors) (ES, 0.76; 95% CI, 0.30–1.21; $P = .001$; 5 studies), with ES in individual studies ranging from -0.31 to 1.05 for process behaviors (one study with a negative ES) and 0.37 to 1.02 for time behaviors. Ten studies evaluated outcomes using direct effects of SBT on patient-related outcomes such as procedural success (eg, cecal intubation rate in colonoscopy) or major complications. A meta-analysis of these studies showed a small pooled ES of 0.45 (95% CI, 0.17–0.72; $P = .001$) and highly consistent results across studies (I^2, 0%).

To examine the stability of these associations and to explore sources of observed heterogeneity further, we performed a preplanned subgroup analysis. Findings in these analyses largely paralleled the main analysis for all outcomes; we show results for process skills (Figure 3*A*) and process behaviors (Figure 3*B*). Notably, both ex vivo animal models and virtual reality systems were associated with favorable effects. Sensitivity analyses excluding nonrandomized studies (process skills: ES, 1.24; 95% CI,

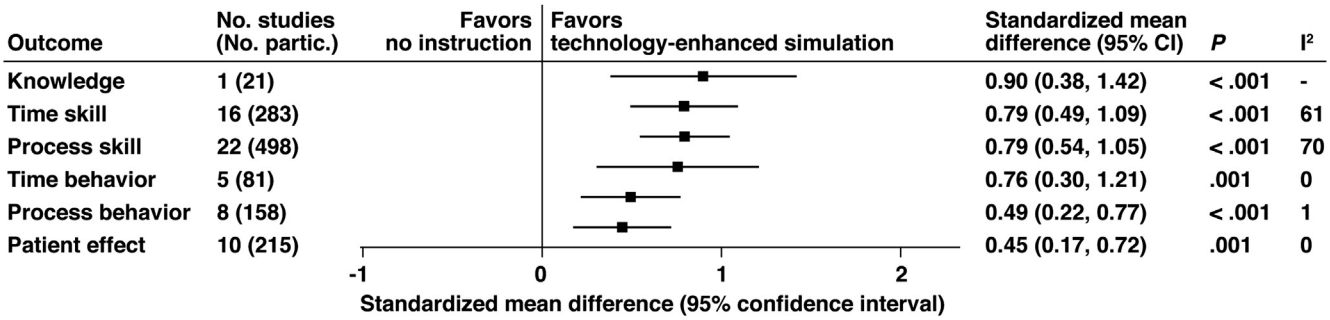

Figure 2. Meta-analysis of studies comparing SBT in gastrointestinal endoscopy with no intervention. Positive standardized mean differences favor the simulation intervention. Results reflect pooling using random-effects meta-analysis. Time outcomes refers to time to completion of procedure in a test setting (time) and in clinical practice (time behavior); process outcomes refers to process of endoscopy (eg, economy of movement, minor errors, or global ratings) in a test setting (process) or in clinical practice (process behavior). Patient effects refers to direct effects on patients (procedure completion rates, major complications, or patient discomfort).

0.93–1.55; 9 studies; process behaviors: ES, 0.63; 95% CI, 0.25–1.10; 5 studies), and studies with imprecise ES (process skills: ES, 0.95; 95% CI, 0.53–1.37; 14 studies; process behaviors: ES, 0.41; 95% CI, 0.12–0.71; 7 studies) likewise yielded similar pooled results. The funnel plot was slightly asymmetric for time and process skills and for patient effects. Assuming this asymmetry reflects publication bias, trim-and-fill analyses showed slightly lower but still statistically significant pooled ES (data not shown).

Three studies reported a negative ES (ie, outcomes were apparently worse after SBT). In the first study, novice gastroenterology fellows who trained using a virtual-reality model had higher patient discomfort scores when subsequently performing procedures on real patients than did a control group without training. The investigators speculated that this might be owing to inaccurate modeling of the actual task (eg, length of scope required to complete the procedure) and inadequate modeling of patient discomfort during training.[24] Another study found slower time skills but higher process skills (eg, mucosa visualized) after 5 simulated colonoscopies without formal training or structured feedback from the simulator.[25] In the third study,

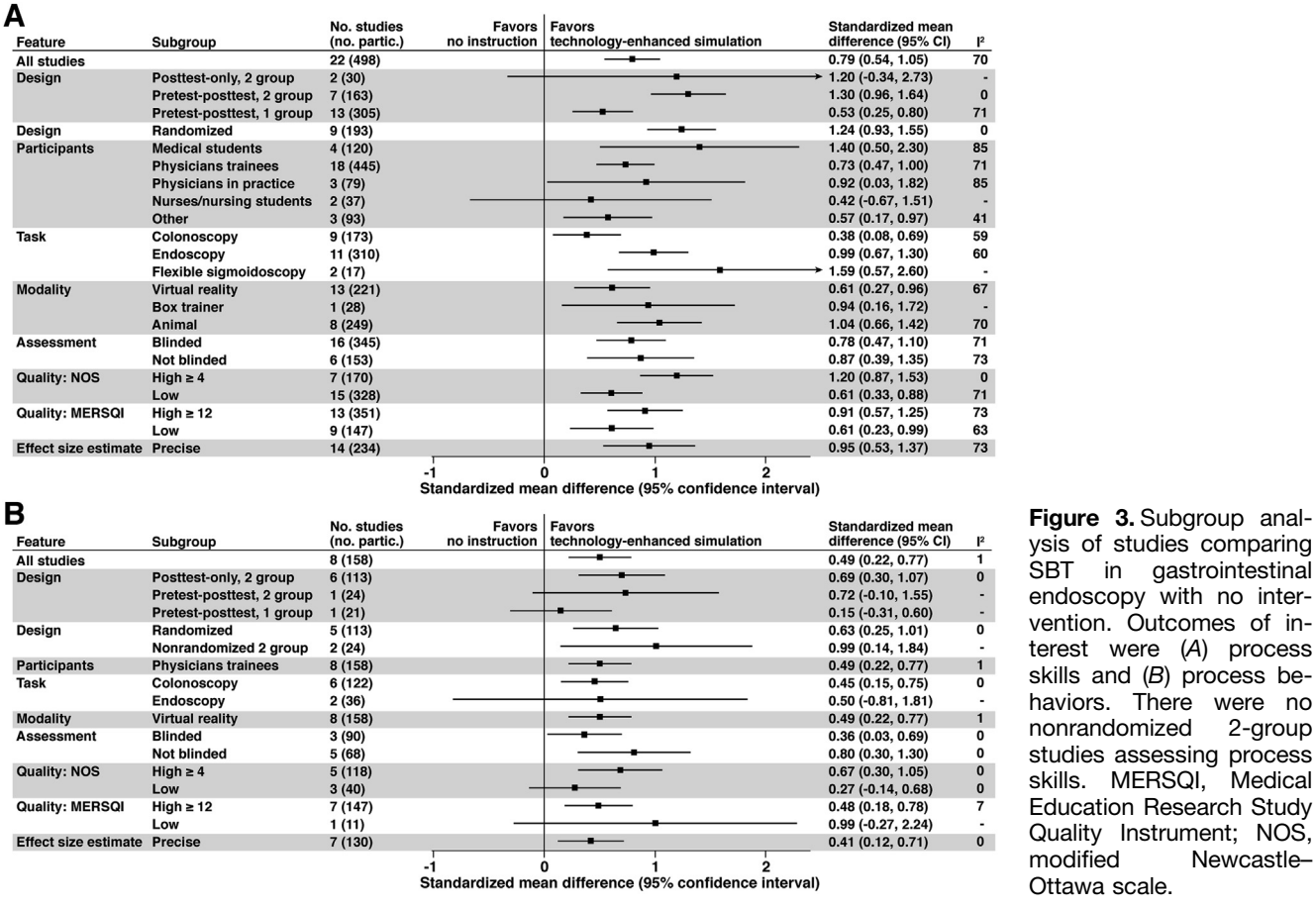

Figure 3. Subgroup analysis of studies comparing SBT in gastrointestinal endoscopy with no intervention. Outcomes of interest were (*A*) process skills and (*B*) process behaviors. There were no nonrandomized 2-group studies assessing process skills. MERSQI, Medical Education Research Study Quality Instrument; NOS, modified Newcastle–Ottawa scale.

intensive colonoscopy training over 5 days was associated with a significant improvement in time skills and process behaviors but with a small, nonsignificant decrease in process skills (simulation-based mucosal visualization decreased from 96.1% at baseline to 95.9% post-training).[26]

Findings from more recent studies identified through an updated search through December 2013 also were similar. In these studies, compared with no intervention, SBT was associated with improvement in trainee skills and behaviors as well as patient outcomes, including in SBT in ERCP (Supplementary Results).

Effectiveness in comparison with nonsimulation training. Two studies compared SBT with nonsimulation training. In an early, unblinded study, trainees randomized to computer-based colonoscopy simulator (with virtual feedback) were less likely to independently successfully complete flexible sigmoidoscopy in real patients (process behavior) as compared with trainees in an apprenticeship model (10 procedures on a real patient followed by expert feedback), although there was no significant difference in time taken to complete sigmoidoscopy (time behavior) or patient tolerance.[27] A more recent assessor-blinded randomized controlled trial compared 16 hours of virtual reality SBT with 16 hours of patient-based training, and reported a significant improvement in skills (in a test setting) and behaviors (in clinical practice) in both groups, as compared with pretraining. The SBT group performed significantly better on simulated cases (process and time skills), and both groups performed similarly well in the context of real patient care (process and time behaviors).[28]

A more recent, 3-arm, randomized controlled trial identified through an updated search, compared the effect of SBT, traditional clinical training, and a combination of the two in enhancing trainee skills in EGD. The investigators observed that all 3 modalities resulted in improving trainee skills, and trainees randomized to SBT or traditional clinical training had comparable performance in a test setting. However, the overall performance (as assessed by a blinded evaluator) was superior for trainees who received a combination of SBT and traditional clinical training, as compared with trainees who received only SBT (Supplementary Results).

Lessons learned from comparisons of two simulation interventions. For evidence-based education to be optimally designed, teachers need guidance on key decisions in course development such as feedback (how and how much to deliver), distribution of practice (1 day or spread over several days), and authenticity (need for fidelity or accurate clinical context).[22] We sought evidence for these and other instructional design features in an inductive analysis of all studies making a comparison with other simulation-based instruction (ie, comparative effectiveness). We found 5 studies making direct comparisons between different simulation-based approaches. As detailed later, one of these studies evaluated training for endoscopy and laparoscopy; the other 4 studies compared different approaches for endoscopy training, and thus offered evidence to inform instructional design.

Two studies addressed the format and timing of feedback in SBT for gastrointestinal endoscopy. One study, using a virtual reality intervention, evaluated the role of instructor feedback. In this study, structured feedback from experts was more effective than virtual feedback in improving process and time skills (ES, 0.58 for both outcomes) in colonoscopy.[29] In another randomized study using a bench-top colonoscopy simulator, the investigators observed that terminal feedback (feedback after the simulation experience) was superior to concurrent feedback (feedback while actively performing simulation) when assessed on a new task (transfer task) (process skills: ES, 0.26; time skills: ES, 0.51).[30]

Two comparative effectiveness studies compared different simulation modalities, which, owing to unique features of each intervention, have somewhat limited generalizability. One study compared virtual reality with an ex vivo cadaveric model using an outcome of satisfaction, and found that 67% of experienced endoscopists and 45% of young trainees preferred the cadaveric model.[23] However, this study did not assess skills or behaviors. Another study compared 2 variants of ex vivo models, and suggested that the newer, compact, lighter version, Compact Erlangen Active Simulator for Interventional Endoscopy (compactEASIE; Erlangen, Germany) was superior to the older Erlanger Endo-Trainer in helping with conceptual understanding.[31] An additional comparative effectiveness study identified through an updated search reported no significant difference between colonoscopy skills acquired through training on high-fidelity, virtual-reality simulation and a low-fidelity, part-task trainer (Supplementary Results).

Finally, one cross-over study sought to determine how well skills gained in training for endoscopy transferred to a laparoscopic surgery task, and vice versa. The investigators found that when medical students were assessed using the same task they had been trained on (eg, trained and assessed on colonoscopy), they performed faster (time skills: ES, 0.33) than when they were assessed on a different task (eg, trained on laparoscopy, assessed on colonoscopy). They interpreted this as supporting the need for task-specific training (vs generalizable psychomotor skills).[32]

Discussion

In this systematic review of 39 studies, nearly 1200 participants at various levels of training, and various SBT modalities of EGD, colonoscopy, and ERCP, we observed that SBT is consistently and significantly better than no intervention in improving trainees' performance as measured both in a test setting (skills assessment) and in clinical practice (behaviors and patient effects). These effects were stable across gastrointestinal endoscopic

procedures, modality of SBT, and training level of participants. SBT also appears to be comparable with traditional patient-based teaching in skills development although the number of studies is few. Studies evaluating the comparative effectiveness of different simulation-based approaches are few, but tentatively suggest that formal expert feedback may be superior to automated virtual feedback, and that terminal feedback may be more effective than concurrent feedback for transfer to new tasks.

Strengths and Limitations

The strengths of this review include the following: (1) comprehensive and systematic literature search with broad inclusion criteria and no restrictions for study design or language; (2) reproducible and rigorous coding; (3) separate comparisons according to outcome and comparator; (4) evaluation of study quality; (5) subgroup analyses to evaluate the stability of findings and to identify potential factors responsible for inconsistencies.

There also were several limitations. First, our search was conducted in May 2011; studies published after that date were not included in the quantitative synthesis. However, we conducted an updated search in December 2013 and have summarized key findings from these studies. Findings from these recent studies also are similar to results from this quantitative synthesis (Supplementary Results). Second, we observed significant clinical and methodologic heterogeneity in the overall analysis. However, this heterogeneity was seen primarily in the strength of the association between SBT and relevant outcomes, and not in the direction of association. We suspect this was in large part owing to variation in SBT modalities and instructional events, outcome measures, and, to a lesser extent, trainees. Although this diversity is a weakness in terms of between-study heterogeneity, it is also a strength in terms of comprehensiveness and breadth of scope. Third, our results were limited by the quality of original studies. Approximately half the studies were nonrandomized, including several single-group, pre-intervention and postintervention studies. Follow-up evaluation was suboptimal in approximately one third of studies and outcome assessment was not blinded (and hence, at risk for assessment bias) in 30%. Fourth, the number of comparative-effectiveness studies (comparing one simulation-based modality with another simulation-based modality) was limited, precluding a detailed analysis. Finally, several endoscopic tasks were represented poorly among the studies identified. For example, we found only 2 studies of training for ERCP (published in 2011 and 2013; Supplementary Table 1) and no studies of training in endoscopic ultrasound. Additional studies on SBT in ERCP and endoscopic ultrasound are warranted.

Comparison With Previous Reviews

One previous review on this topic included 5 studies of simulation-based endoscopy training,[3] and a recent Cochrane review identified 13 studies of virtual reality endoscopy training.[5] To these narrative reviews we add 26 additional studies and a quantitative synthesis of evidence confirming moderate to large effects on multiple training-related outcomes of SBT, across various subgroups. Our conclusions regarding the efficacy of SBT in gastrointestinal endoscopy are similar to those for SBT in training in other endoscopic procedures including bronchoscopy, urologic endoscopy, and laparoscopic surgery,[33–35] and simulation in general.[6,36]

The paucity of comparative effectiveness studies limits the conclusions we can draw regarding principles of effective instructional design and selection of modalities. However, previous reviews have explored these issues in simulation broadly,[19,22,37,38] and have identified several best practices including feedback, mastery learning, repetitive practice, and range of difficulty. For example, a recent systematic review found strong evidence that instructor-generated feedback was superior to simulator-generated feedback for simulation-based procedural skills training (including endoscopy and several other procedural tasks).[37] The timing of the feedback also is important. In contrast to a recent review of laparoscopic surgical skills that found that concurrent feedback appears to be superior to terminal feedback in short-term skills performance of trainees,[35] the one relevant study in this review found terminal feedback to be superior when performing a new task (skill transfer). Further research on the timing and format of feedback would extend our understanding of effective simulation instructional design.

Implications and Conclusions

Our findings have important implications for current practice and future research. Simulation-based education in diagnostic and therapeutic gastrointestinal endoscopy training is effective, and this learning transfers to patient care. This lends support to recent American Council for Graduate Medical Education training guidelines, which mandate the incorporation of SBT in all gastroenterology fellowship programs.[2] SBT may be particularly useful in early clinical endoscopy training by providing a low-stress learning environment with no risk for either trainees or patients. For basic endoscopy skills, the current evidence suggests little difference across simulation modalities. Training for more advanced techniques (such as loop reduction, therapeutic interventions, and ERCP) may require greater visual and haptic fidelity, and requires further study.

All but 1 of 32 studies making a comparison with no intervention found benefit in at least 1 outcome. Given the uniformity of effect, we propose that the field is

unlikely to profit from further studies using single-group, pre-post designs and comparisons with no-intervention or placebo-intervention controls. We instead encourage investigators to pursue studies using carefully conceptualized active comparison groups (comparative effectiveness studies).[39] Such studies will be most useful when the comparison can be generalized to a broad range of future applications.

The 2 studies showing a decrease in performance after training[24,25] suffered from deficiencies in instructional design (lack of accurate modeling of the clinical task and absence of effective instructional design). These studies highlight the need for careful attention to the instructional design[19] and functional task alignment[40] when planning and implementing SBT. Additional research is warranted to identify the ideal simulation modalities (virtual-reality or ex vivo models) and instructional methods (optimal timing of SBT in the training continuum, duration of SBT, type of supervision and feedback, and so forth) for specific diagnostic and therapeutic endoscopic tasks. Currently, ex vivo and animal models appear to be more realistic but require greater effort to obtain and maintain than virtual reality and synthetic materials trainers. Although computer-based virtual reality simulators have shown promise, their up-front cost is prohibitive for many programs, and moreover some of the purported enhancements (eg, visual fidelity and haptic force feedback) have little evidence to support their added value.[35,40] Comparative-effectiveness studies comparing different simulation approaches will contribute greatly to this evidence base.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at http://dx.doi.org/10.1016/j.cgh.2014.01.037.

References

1. Sedlack RE. The state of simulation in endoscopy education: continuing to advance toward our goals. Gastroenterology 2013;144:9–12.
2. ACGME program requirements for graduate medical education in gastroenterology (internal medicine) section IV.A.6.b. Available at: http://www.acgme.org/acgmeweb/Portals/0/PFAssets/2013-PR-FAQ-PIF/144_gastroenterology_int_med_07132013.pdf. Accessed June 30, 2013.
3. Sturm LP, Windsor JA, Cosman PH, et al. A systematic review of skills transfer after surgical simulation training. Ann Surg 2008;248:166–179.
4. Gerson LB. Evidence-based assessment of endoscopic simulators for training. Gastrointest Endosc Clin N Am 2006;16:489–509.
5. Walsh CM, Sherlock ME, Ling SC, et al. Virtual reality simulation training for health professions trainees in gastrointestinal endoscopy. Cochrane Database Syst Rev 2012;6:CD008237.
6. Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011;306:978–988.
7. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009;151:264–269.
8. Desilets DJ, Banerjee S, Barth BA, et al. Endoscopic simulators. Gastrointest Endosc 2011;73:861–867.
9. Baillie J. Endoscopic retrograde cholangiopancreatography simulation. Gastrointest Endosc Clin N Am 2006;16:529–542.
10. Cohen J. The first International Conference on Endoscopy Simulation: consensus statement. Gastrointest Endosc Clin N Am 2006;16:583–591.
11. Kefalides PT, Gress F. Simulator training for endoscopic ultrasound. Gastrointest Endosc Clin N Am 2006;16:543–552.
12. Matthes K. Simulator training in endoscopic hemostasis. Gastrointest Endosc Clin N Am 2006;16:511–527.
13. Sedlack RE. Simulators in training: defining the optimal role for various simulation models in the training environment. Gastrointest Endosc Clin N Am 2006;16:553–563.
14. Williams CB, Thomas-Gibson S. Rational colonoscopy, realistic simulation, and accelerated teaching. Gastrointest Endosc Clin N Am 2006;16:457–470.
15. Sutherland LM, Middleton PF, Anthony A, et al. Surgical simulation: a systematic review. Ann Surg 2006;243:291–300.
16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.
17. Reed DA, Cook DA, Beckman TJ, et al. Association between funding and quality of published medical education research. JAMA 2007;298:1002–1009.
18. Cook DA, Levinson AJ, Garside S, et al. Internet-based learning in the health professions: a meta-analysis. JAMA 2008;300:1181–1196.
19. Cook DA, Hamstra SJ, Brydges R, et al. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. Med Teach 2013;35:e867–e898.
20. Higgins JPT, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–560.
21. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum, 1988.
22. Issenberg SB, McGaghie WC, Petrusa ER, et al. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach 2005;27:10–28.
23. Kiesslich R, Moenk S, Reinhardt K, et al. Combined simulation training: a new concept and workshop is useful for crisis management in gastrointestinal endoscopy [German]. Z Gastroenterol 2005;43:1031–1039.
24. Sedlack RE. Validation of computer simulation training for esophagogastroduodenoscopy: pilot study. J Gastroenterol Hepatol 2007;22:1214–1219.
25. 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.
26. Thomas-Gibson S, Bassett P, Suzuki N, et al. Intensive training over 5 days improves colonoscopy skills long-term. Endoscopy 2007;39:818–824.
27. 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.

28. 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.
29. Kruglikova I, Grantcharov TP, Drewes AM, et al. 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.
30. Walsh CM, Ling SC, Wang CS, et al. Concurrent versus terminal feedback: it may be better to wait. Acad Med 2009; 84(Suppl):S54–S57.
31. Hochberger J, Euler K, Naegel A, et al. The compact Erlangen Active Simulator for interventional endoscopy: a prospective comparison in structured team-training courses on 'endoscopic hemostasis' for doctors and nurses. Scand J Gastroenterol 2004;39:895–902.
32. Buzink SN, Goossens RHM, Schoon EJ, et al. Do basic psychomotor skills transfer between different image-based procedures? World J Surg 2010;34:933–940.
33. Kennedy CC, Maldonado F, Cook DA. Simulation-based bronchoscopy training: systematic review and meta-analysis. Chest 2013;144:183–192.
34. Skolarikos A, Gravas S, Laguna MP, et al. Training in ureteroscopy: a critical appraisal of the literature. BJU Int 2011;108: 798–805.
35. Zendejas B, Brydges R, Hamstra SJ, et al. State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. Ann Surg 2013;257:586–593.
36. Cook DA, Brydges R, Hamstra SJ, et al. Comparative effectiveness of technology-enhanced simulation versus other instructional methods: a systematic review and meta-analysis. Simul Healthc 2012;7:308–320.
37. Hatala R, Cook DA, Zendejas B, et al. Feedback for simulation-based procedural skills training: a meta-analysis and critical narrative synthesis. Adv Health Sci Educ Theory Pract 2014; 19:251–272.
38. Cook DA, Brydges R, Zendejas B, et al. Mastery learning for health professionals using technology-enhanced simulation: a systematic review and meta-analysis. Acad Med 2013; 88:1178–1186.
39. Cook DA. If you teach them, they will learn: why medical education needs comparative effectiveness research. Adv Health Sci Educ Theory Pract 2012;17:305–310.
40. Hamstra SJ, Brydges R, Hatala R, et al. Reconsidering fidelity in simulation-based training: a critique based on evidence from a systematic review. Acad Med 2014;89:387–392.
41. Tuggy ML. Virtual reality flexible sigmoidoscopy simulator training: impact on resident performance. J Am Board Fam Pract 1998;11:426–433.
42. Ferlitsch A, Glauninger P, Gupper A, et al. Evaluation of a virtual endoscopy simulator for training in gastrointestinal endoscopy. Endoscopy 2002;34:698–702.
43. Kneebone RL, Nestel D, Moorthy K, et al. Learning the skills of flexible sigmoidoscopy–the wider perspective. Med Educ 2003; 37(Suppl 1):50–58.
44. Neumann M, Hahn C, Horbach T, et al. Score card endoscopy: a multicenter study to evaluate learning curves in 1-week courses using the Erlangen Endo-Trainer. Endoscopy 2003;35:515–520.
45. Neumann M, Siebert T, Rausch J, et al. Scorecard endoscopy: a pilot study to assess basic skills in trainees for upper gastrointestinal endoscopy. Langenbeck Arch Surg 2003;387:386–391.
46. Di Giulio E, 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.
47. 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.
48. Sedlack RE, Kolars JC, Alexander JA. Computer simulation training enhances patient comfort during endoscopy. Clin Gastroenterol Hepatol 2004;2:348–352.
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. Ahlberg G, Hultcrantz R, Jaramillo E, et al. Virtual reality colonoscopy simulation: a compulsory practice for the future colonoscopist? Endoscopy 2005;37:1198–1204.
51. Hochberger J, Matthes K, Maiss J, et al. Training with the compactEASIE biologic endoscopy simulator significantly improves hemostatic technical skill of gastroenterology fellows: a randomized controlled comparison. Gastrointest Endosc 2005; 61:204–215.
52. Matthes K, Cohen J, Kochman ML, et al. Efficacy and costs of a one-day hands-on EASIE endoscopy simulator train-the-trainer workshop. Gastrointest Endosc 2005;62:921–927.
53. 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.
54. 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.
55. Maiss J, Prat F, Wiesnet J, et al. The complementary Erlangen active simulator for interventional endoscopy training is superior to solely clinical education in endoscopic hemostasis–the French training project. Eur J Gastroenterol Hepatol 2006;18: 1217–1225.
56. Ritter EM, McClusky DA III, Gallagher AG, et al. Perceptual, visuospatial, and psychomotor abilities correlate with duration of training required on a virtual-reality flexible endoscopy simulator. Am J Surg 2006;192:379–384.
57. Thomson M, Heuschkel R, Donaldson N, et al. Acquisition of competence in paediatric ileocolonoscopy with virtual endoscopy training. J Pediatr Gastroenterol Nutr 2006;43:699–701.
58. 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.
59. Maiss J, Millermann L, Heinemann K, et al. The compactEASIE® is a feasible training model for endoscopic novices: a prospective randomised trial. Dig Liver Dis 2007;39:70–78.
60. 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.
61. Yi SY, Ryu KH, Woo HS, et al. Quantitative analysis of colonoscopy skills using the KAIST-Ewha Colonoscopy Simulator II. Frontiers in the Convergence of Bioscience and Information Technologies 2007; http://dx.doi.org/10.1109/FBIT.2007.58.
62. Shirai Y, Yoshida T, Shiraishi R, et al. Prospective randomized study on the use of a computer-based endoscopic simulator for

training in esophagogastroduodenoscopy. J Gastroenterol Hepatol 2008;23:1046–1050.
63. Yi SY, Ryu KH, Na YJ, et al. Improvement of colonoscopy skills through simulation-based training. Stud Health Technol Inform 2008;132:565–567.
64. Haycock AV, Youd P, Bassett P, et al. Simulator training improves practical skills in therapeutic GI endoscopy: results from a randomized, blinded, controlled study. Gastrointest Endosc 2009;70:835–845.
65. Yi SY, Ryu KH, Woo HS, et al. Sectional analysis of learning on the KAIST-Ewha Colonoscopy Simulation II. Stud Health Technol Inform 2009;142:432–434.
66. 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.
67. Haycock AV, East JE, Swain D, et al. Development of a novel esophageal stricture simulation. Dig Dis Sci 2010;55:321–327.
68. Kruglikova I, Grantcharov TP, Drewes AM, et al. Assessment of early learning curves among nurses and physicians using a high-fidelity virtual-reality colonoscopy simulator. Surg Endosc 2010;24:366–370.
69. Luursema JM, Buzink SN, Verwey WB, et al. Visuo-spatial ability in colonoscopy simulator training. Adv Health Sci Educ Theory Pract 2010;15:685–694.

Reprint requests
Address requests for reprints to: David A. Cook, MD, MHPE, Division of General Internal Medicine, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, Minnesota 55905. e-mail: cook.david33@mayo.edu.

Acknowledgments
The authors thank Ryan Brydges, PhD; Patricia J. Erwin, MLS; Stanley J. Hamstra, PhD; Rose Hatala, MD; Jason H. Szostek, MD; Amy T. Wang, MD; and Benjamin Zendejas, MD, for their assistance in the literature search and initial data acquisition.

Conflicts of interest
The authors disclose no conflicts.

Funding
Supported by intramural funds, including an award from the Division of General Internal Medicine, Mayo Clinic.

The funding sources for this study played no role in the design and conduct of the study; in the collection, management, analysis, and interpretation of the data; or in the preparation of the manuscript. The funding sources did not review the manuscript.

Supplementary Results

Key Findings From Additional Studies

In 6 studies comparing SBT with no intervention identified through an updated search, findings were similar to those observed in previous studies. SBT was associated with superior trainee performance in a test setting as well as in clinical practice, as compared with no intervention.[1–6] The rate of independent cecal intubation also was higher in trainees who underwent SBT in colonoscopy,[2,4] although a significant difference in the amount of sedative medication use or use of assistive maneuvers such as patient repositioning was not identified.[2] Two recent studies also assessed the efficacy of SBT in ERCP training using a mechanical simulator in addition to standard clinical training.[3,5] In both studies, trainees randomized to SBT had higher rates of independent biliary cannulation, in a shorter amount of time. In one study, through indirect comparison, investigators observed no additional benefit of multiple unsupervised SBT sessions in ERCP, after the initial supervised SBT.[5]

In a more recent 3-arm randomized controlled trial of 28 medical students comparing SBT, traditional clinical training, and a combination of the two, for training in EGD, Ende et al[7] observed that SBT was comparable with traditional clinical training in enhancing trainee skills in a test setting. However, the overall performance (as assessed by a blinded evaluator) was superior for trainees who received a combination of SBT and traditional clinical training, as compared with trainees who received only SBT.[7] In another RCT comparing a high-fidelity, virtual reality simulator with a low-fidelity, part-task trainer for SBT in colonoscopy, Ahad et al[8] observed that both resulted in improvement in trainee performance and there was no significant difference between skills acquired with the low- and high-fidelity SBT.

References

1. Gotzberger M, Rosch T, Geisenhof S, et al. Effectiveness of a novel endoscopy training concept. Endoscopy 2011;43:802–807.
2. 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.
3. 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.
4. Bai Y, Zhi FC, Liu SD, et al. [Control study on colonoscopy skills acquiring from endoscopic simulation system transferring to patients]. Zhonghua Yi Xue Za Zhi 2012;92:1285–1287.
5. Liao WC, Leung JW, Wang HP, et al. Coached practice using ERCP mechanical simulator improves trainees' ERCP performance: a randomized controlled trial. Endoscopy 2013;45:799–805.
6. Telem DA, Rattner DW, Gee DW. Endoscopic simulator curriculum improves colonoscopy performance in novice surgical interns as demonstrated in a swine model. Surg Endosc 2014;28:1494–1499.
7. 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.
8. Ahad S, Boehler M, Schwind CJ, et al. The effect of model fidelity on colonoscopic skills acquisition. A randomized controlled study. J Surg Educ 2013;70:522–527.

Supplementary Table 1. Key Results From Additional Studies Identified on Screening Articles From May 2011 to December 2013

Study	Task	Study design	Trainees; N	Intervention/ comparison	Outcomes	Key results
Comparison of SBT vs no intervention						
Gotzberger et al, 2011[1]	EGD and colonoscopy	1-group, pre- and postintervention study	Residents and expert endoscopists; 78	Intervention: GATE, with theoretical training, followed by supervised and unsupervised virtual-reality SBT	Knowledge; skills (process and time) in a virtual-reality simulator	The GATE program was associated with improved theoretical knowledge, compared with pretest performance (pre- vs postintervention score: 1.69 $\pm$ 1.01 vs 3.27 $\pm$ 1.30 points; $P < .01$) SBT was associated with an improvement in practical skills (pre- vs postintervention score in terms of difficulty of snare polypectomy–higher score worse: 3.05 $\pm$ 0.65 vs 2.52 $\pm$ 0.59; $P = .08$), as well as time to procedure completion (pre- vs postintervention time for snare polypectomy: 445 $\pm$ 189 s vs 274 $\pm$ 129 s; $P < .01$)
Kaltenbach et al, 2011[2]	Colonoscopy	1-group, pre- and postintervention study	Gastroenterology trainees not proficient in colonoscopy; 3	Intervention: supervised virtual-reality SBT using a colon model with 3D instrument visualization	Behaviors (process and time); patient outcomes	SBT was associated with improved cecal intubation rate (pre- vs postintervention: 43% vs 76%), time (pre- vs postintervention: 17.7 $\pm$ 11.3 vs 13.5 $\pm$ 6.5 min) and performance (overall insertion skill score, pre- vs postintervention: 4.4 $\pm$ 2.4 vs 5.9 $\pm$ 2.4; $P < .01$) as assessed by a blinded evaluator SBT was not associated with a significant change in use of sedative medications and patient repositioning
Lim et al, 2011[3]	ERCP	2-group, RCT, comparing SBT with no intervention	Gastroenterology fellows with <30 ERCPs at time of entry into study; 16	Intervention: 2 supervised sessions on ERCP mechanical simulator Comparison: no intervention (both groups received standard clinical training)	Behaviors (process and time)	Trainees randomized to SBT were more likely to achieve successful independent biliary cannulation (intervention vs comparison: 69.6% vs 47.1%; $P = .02$), in shorter time (mean cannulation time, intervention vs comparison: 4.7 $\pm$ 4.2 vs 10.3 $\pm$ 14.1 min; $P < .01$), as compared with trainees who did not receive any intervention There was no difference in the overall trainee competence scores between the 2 groups (trainee competency score on 5-point scale–higher score better, intervention vs comparison: 4.2 $\pm$ 0.9 vs 4.1 $\pm$ 0.8; $P = .36$), as assessed by a blinded evaluator
Bai et al, 2012[4]	Colonoscopy	2-group, nonrandomized trial, comparing SBT with no intervention	Trainees; 8	Intervention: virtual-reality training using AccuTouch endoscopy simulator (Saint-Laurent, Quebec, Canada) Control: no intervention (both groups received standard clinical training)	Patient outcomes (procedural completion)	SBT was associated with increased completion of procedure (number of procedures completed independently, intervention vs comparison: 21 vs 2; $P < .01$), and "decreased incompetence"[4]

Supplementary Table 1. Continued

Study	Task	Study design	Trainees; N	Intervention/ comparison	Outcomes	Key results
Liao et al, 2013[5]	ERCP	2-group RCT comparing SBT with no intervention	Advanced gastroenterology fellows; 16	Intervention: single supervised, with or without multiple uncoached practice on an ERCP mechanical simulator Comparison: no intervention (both groups received standard clinical training)	Behaviors (process); patient outcomes (procedural completion)	Trainees randomized to SBT were more likely to achieve successful biliary cannulation (intervention vs comparison: 72.4% vs 46.7%; $P < .01$) and had a better overall performance score (intervention vs comparison: 3.9 $\pm$ 0.8 vs 3.2 $\pm$ 0.7; $P =$ not reported) as assessed by blinded evaluators, compared with patients who did not receive any intervention Through indirect comparison there was no additional benefit of multiple unsupervised SBT in ERCP, after initial supervised SBT
Telem et al, 2013[6]	Colonoscopy	1-group, pre- and postintervention study	Postgraduate trainees (surgical interns); 9	Intervention: 12 h of independent training on a virtual reality trainer (GI Mentor; Simbionix, Cleveland, OH)	Skills (process and time) in a porcine model	SBT was associated with improved cecal intubation rate (pre- vs postintervention: 11% vs 67%; $P <$ .05) There were significant improvements in the Global Assessment of Endoscopic Skills performance (pre- vs postintervention score: 11.4 $\pm$ 1.8 vs 14.6 $\pm$ 1.3; $P < .01$), as assessed by a blinded evaluator
Comparison of SBT vs nonsimulation intervention (and vs no intervention)						
Ende et al, 2012[7]	EGD	3-group RCT comparing SBT with nonsimulation intervention and no intervention	Medical residents with no endoscopy experience; 28	Group 1: traditional clinical training + SBT Group 2: traditional clinical training alone Group 3: SBT alone; SBT involved virtual-reality simulator with 10 sessions over 4 months	Skills (process and time) in a virtual-reality simulator	All groups had improvement in process (pre- vs postintervention scores in groups 1, 2, and 3 on a 10-point scale–higher score better: 2.3 vs 4.3; 2.2 vs 3.4; 2.7 vs 4.8, respectively; all $P < .05$) and time skills (pre- vs postintervention scores in groups 1, 2, and 3: 195 vs 119 s; 261 vs 150 s; 165 vs 117 s, respectively; P for groups 1 and 2 $< .05$, P for group 3 $= .07$) after clinical and/or SBT, as compared with pretraining performance The overall assessment of procedural performance was superior for group 1 compared with group 3 (group 1 vs group 3, performance score on 10-point scale–6.6 vs 5.1; $P = .03$) as assessed by a blinded evaluator, although there was no difference in time to procedure completion between groups (group 1 vs group 3: 822 $\pm$ 163 vs 968 $\pm$ 139, $P =$ .20); there were no differences in process and time skills between groups 1 and 2, and groups 2 and 3

Comparison of one SBT modality with another						
Ahad et al, 2013[8]	Colonoscopy	2-group RCT comparing 2 different SBT modalities	Medical students (third- and fourth-year); 32	Intervention: high-fidelity, virtual-reality simulator Comparison: low-fidelity, part-task trainer	Skills (process and time) in a virtual-reality simulator	SBT with both high- and low-fidelity simulators resulted in decrease in insertion time (pre- vs postintervention with high-fidelity and low-fidelity simulator: 426 $\pm$ 169 vs 331 $\pm$ 219 s, $P = .02$; 481 $\pm$ 236 vs 255 $\pm$ 106 s, $P < .01$, respectively) and increased total mucosal visualization (percentage of total mucosa visualized, pre- vs postintervention with high-fidelity and low-fidelity simulator: 86% vs 96%, $P = .03$; 88% vs 97%, $P = .04$, respectively) There was no significant difference between skills acquired with the low- and high-fidelity virtual-reality based training

GATE, gastroenterologic education-training endoscopy; RCT, randomized controlled trial; 3D, 3-dimensional.