*P*reservation and *I*ncorporation of *V*aluable Endoscopic *I*nnovations (PIVI)

The American Society for Gastrointestinal Endoscopy PIVI on the use of endoscopy simulators for training and assessing skill (Long form)

The PIVI Initiative

The PIVI initiative is an ASGE program that aims to identify important clinical questions related to endoscopy and to establish a priori diagnostic and/or therapeutic thresholds for endoscopic technologies designed to resolve these clinical questions. Additionally, PIVIs may also outline the data and or the research study design required for proving an established threshold is met. Once endoscopic technologies meet an established PIVI threshold, those technologies are appropriate to incorporate into clinical practice presuming the appropriate training in that endoscopic technology has been achieved. The ASGE encourages and supports the appropriate use of technologies that meet its established PIVI thresholds.

The PIVI initiative was developed primarily to direct endoscopic technology development toward resolving important clinical issues in endoscopy. The PIVI initiative is also designed to minimize the possibility that potentially valuable innovations are prematurely abandoned due to lack of utilization and to avoid widespread use of an endoscopic technology before clinical studies documenting their effectiveness have been performed. The following document, or PIVI, is one of a series of statements defining the diagnostic or therapeutic threshold that must be met for a technique or device to become considered appropriate for incorporation into clinical practice. It is also meant to serve as a guide for researchers or those seeking to develop technologies that are designed to improve digestive health outcomes.

An ad hoc committee under the auspices of the existing ASGE Technology and Standards of Practice Committees Chairs develops PIVIs. An expert in the subject area chairs the PIVI committee, with additional committee members chosen for their individual expertise. In preparing this document, evidence-based methodology was employed, using a MEDLINE and PubMed literature search to identify pertinent clinical studies on the topic. PIVIs are ultimately submitted to the ASGE Governing Board for approval, as is done for all Technology and Standards of Practice documents. This document is provided solely for educational and informational purposes and to support incorporating these endoscopic technologies into clinical practice. It should not be construed as establishing a legal standard of care.

See published PIVI summary in *GIE: Gastrointestinal Endoscopy*
Use of Endoscopy Simulator for Training and Assessing Skill
Gastrointest Endosc 2012:76:471-475

Use of Endoscopy Simulators for Training and Assessing Skill
PIVI
(Long form)

I. General clinical area of this PIVI

This PIVI will review the current literature on simulator use in endoscopy and assess what data is required to support a wider adoption of their use for endoscopy training and skills assessment.

II. Specific clinical issues addressed by this PIVI

a. How much benefit must be demonstrated from the use of simulators to justify widespread adoption into standard endoscopy training?

b. How reliable do simulator-based assessments need to be as a predictor of patient-based skills to justify their use in credentialing and re-credentialing for endoscopy?

Definitions and assumptions:

For the purpose of this PIVI, the term "simulator" refers to all educational tools which allow for repetitive instruction in an environment free of any actual patient care delivery. This includes static models, *ex vivo* animal tissue models, live animal models, and computer simulators.

Comparisons of simulator-based education to standard methods alone can only be made once learning curves are established for standard instruction using objective measures that encompass technical and cognitive skill components of a particular procedure. Skills assessment requires the measurement of agreed-upon parameters and a consensus on what level of performance is acceptable for individuals at the completion of training based on available benchmarks for a given procedure (i.e., minimal competence parameters). In other words, competency must first be defined for each procedure based on objective parameters.

This PIVI does not attempt to define minimal competency parameters for any particular procedure nor delineate the specific parameters for each technique that trainees must master to become competent. However, it does define the expected level of agreement simulation assessment should have to patient-based minimal competence parameters in order to be a useful assessment instrument.

III. Threshold recommended for this PIVI

a. Threshold for incorporation into training

For an endoscopy simulator to be integrated into the standard instruction for a procedure, it must demonstrate a 25% or greater reduction in the median number of clinical cases required for the trainees to achieve the minimal competence parameters for that procedure.

b. Threshold for assessing skill

Simulator-based assessment tools must be procedure-specific and predictive of independently defined minimal competence parameters from real procedures with a kappa value of at least 0.70 for high-stakes assessment.

IV. Summary explanation of thresholds recommended for this PIVI

a. Training

The principle way in which simulators can have a meaningful impact on training would be for them to lead to a significant acceleration of the learning curves to the achievement of competence.[1-3] For colonoscopy, current simulators have demonstrated a benefit in skill acquisition for the first 20-80 cases performed by novices but *no reduction* in the median number of cases required to achieve technical and cognitive competency.[4, 5] With improved realism of models and perhaps more rigorous simulator experience, consensus opinion of the panel of experts was that some modest impact on the learning curve could realistically be achievable. A threshold was chosen which was felt to be both theoretically attainable but also sufficiently high to justify the expense and effort involved in purchasing simulators and incorporating them into the training program. The panel opined that given the expense and effort involved, a reduction in training times or procedure numbers of at least 25% would be required (Figure 1). A more modest 10% benefit was felt to be insufficient to justify the investment in simulation devices by training programs and, based on the results from the existing literature, a 50% reduction in training times/number of cases was felt to be unattainable by any simulator in the near term. While a reduction in the learning curve > 25% is desirable, given both the data on current models and the anticipated expense required to develop simulators that could produce a greater impact on the rate of skill acquisition, current expert consensus arrived at the threshold of 25%.

However, different factors may influence decisions on the part of training programs as to whether to purchase and integrate simulators for particular procedures. These include the cost of simulator purchase and upkeep, the manpower required to utilize it properly, the relative abundance of real cases in a given procedure available for trainee instruction, and the relative complication risk of a given procedure. Other potential advantages to simulation might include rehearsal of the use of new technology, team training and/or review of pathological findings. Regardless of the external pressures to satisfy difficult-to-meet training needs at a particular institution, a measurable benefit from simulator training must be observed and the curriculum (simulator plus traditional proctored human cases) must result in the trainee successfully reaching the minimal level of skill required to perform the procedure independently. An important distinction is to be made between the incorporation of simulators into a training

program for specific purposes and reliance upon them to conduct proper training. The modest enhancement of training shown thus far in the literature for colonoscopy computer simulators underscores the point that the 25% threshold refers to the use of simulators only as a complementary tool to standard proctored endoscopy education by dedicated teachers.

To date, there has been scant evidence that patients materially benefit when trainees doing their procedures have had prior simulator experience. However, future investigation that could demonstrate a meaningful benefit to patients from simulator-augmented curricula, in terms of decreased adverse events, improved satisfaction or better compliance with screening may prompt re-evaluation of the threshold criteria for utilizing simulators.

We must also acknowledge that certain new technologies and procedures may well be introduced with accompanying recommended training programs which integrate some simulator experience. When no standard training program that does not involve the use of simulators for a technique exists, it will be particularly difficult to obtain objective evidence that the simulator work has a significant impact on the development of proficiency.

b. Assessing skill

It is the consensus of this PIVI that strong predictive correlation between simulation performance to actual procedures is necessary for validation and acceptance of a simulator's use in high-stakes assessment.

In education studies, a positive correlation coefficient between two different assessment methods of 0.65 is generally accepted to demonstrate a reasonably strong predictive value; ideally a coefficient of 0.70 or higher is suggested for high-stakes assessment. Correlation coefficients of 0.85 or higher are rarely achieved in education research.[6]As such, a correlation coefficient between simulation scoring of a skill and the scores of the same skill obtained from clinical assessment of patient-based endoscopy would be expected to be at a minimum of 0.65 to be considered useful for assessment of any type, and 0.70 or higher for high-stakes assessment. These goals are obtainable. Surgical literature reports the achievement of similar predictive values in the examinations of proficiency in laparoscopy correlated to competent levels of performance on real procedures.[7,8] In particular, the Fundamentals of Laparoscopic Surgery (FLS) program has been extensively validated and is now a requirement by the American Board of Surgery. Part of the process involved the development of a validated measure of intraoperative skill (GOALS) to use as an outcome measure.[9] The correlation coefficient for GOALS scores during laparoscopic cholecystectomy and performance on the FLS simulator was 0.81.[10] Practice on the FLS simulator was subsequently shown to improve performance in the operating room. Approximately 7.5 hours of simulator practice was needed for first year residents to perform at the level of third year residents in the operating room during laparoscopic cholecystectomy.[11] To date, there are no similar validation studies in the endoscopic literature.

It is expected that performance assessments on simulators will eventually play a role in the high-stakes assessment of competency (i.e., credentialing/certification) and, as such, must have high predictive value for quality performance of the procedure of interest. The demonstrated ability to simply distinguish novice from expert on the model is necessary but not sufficient to guarantee that the evaluation tool will be of clinical utility. The ideal tests will have high interobserver agreement and be geared to finely discriminate trainees at various stages of learning. They must be able to distinguish wide ranges of skill from novice to expert. Global rating systems for procedures that measure both technical and cognitive proficiency will be helpful in allowing for the validation of simulator based skill assessment.

Background on the clinical problems related to this PIVI

a. Training

Since the early days of flexible endoscopy, educators have recognized the potential for simulators to enhance the training of students to gain proficiency. What began with crude static models to provide familiarity with basic dials and scope handling has evolved in the past 15 years into a wide array of *ex vivo* animal tissue and computer virtual reality simulators. The development and capabilities have been well chronicled in the literature, as have many efforts to demonstrate their usefulness particularly in the area of training.[12,13]

The theoretical benefits of simulator training are intuitive. They can provide a student with a relaxed opportunity for repetitive practice of skills including those that might not be encountered with sufficient frequency during the course of a standard training program. Improving basic skills prior to actual patient experience could result in reduced patient discomfort, though this has been documented in only one study.[5, 15] For certain higher risk procedures such as endoscopic retrograde cholangiopancreatography (ERCP), there is the potential for reducing risk to patients undergoing procedures in which novices are participating. Manpower limitations of available endoscopic trainers or cost considerations of increased time trainers must spend away from their clinical duties would support the use of tools that might either shorten the learning curve or allow students to do more of their instruction independently.

While the use of simulators has become much more widespread, particularly via the use of *ex vivo*-based hands on training courses by ASGE at its national training facility at the Institute for Training and Technology Center (ITT) in Oak Brook, IL and at many regional courses throughout the world, there is no consensus to date on just how much of a role they should play in standardized training. The literature summarized in the tables below will define the settings in which simulators have been studied, the endpoints used to assess their utility, and the degree of benefit that has been demonstrated to date.

The question of how good simulators need to be to warrant their use depends on many variables. It begins with a consideration of what are the unmet needs which simulator use might address and a thorough review of their current capabilities and performance. Perhaps the best way to

identify unmet needs within standard endoscopy education is to detect variation in the objective skill level of individuals who have completed a program of training.[15] The burgeoning movement to define quality measures and measure outcome of endoscopy as it is performed has only recently taken hold, and this will be needed to best uncover disparities in competency among recent trainees. Low-volume, high-risk procedures clearly carry a greater incentive for safe repetitive realistic practice than high-volume procedures which offer ample opportunity for skill mastery using standard proctored teaching. Practitioners with less access to sufficient case volume for training might call for simulators to supplant a portion of the real case experience, but this cannot come at the expense of training to accepted benchmarks of independent performance of high quality and safe procedures.

Apart from using models to hasten the learning curve, there are other potential advantages to simulation-like rehearsal of the use of new technology, team training, and review of pathology. Ultimately, the decision about whether to incorporate these technologies into a training program must rely upon data regarding the magnitude of benefits reasonable to expect, the expected resulting cost savings, the initial and ongoing expenses associated with the simulator work, and the local needs of the institution.

b. Assessing skill

The endpoint of endoscopic training is the acquisition of competency to perform procedures independently. Professional societies charged with educating future endoscopists and the public at-large have a vested interest in ensuring that the individuals credentialed to perform endoscopy are able to provide high-quality care. Key to this need for quality assurance is the impetus to move from subjective assessments of trainees' skill to more objective and validated means of doing so before they are credentialed.

Much effort has been devoted, and much more is still required, to define what specific skills are required to become competent in each procedure, to determine minimal standards of proficiency, and to devise ways to objectively assess whether an individual has met that threshold.

Certain guiding principles have evolved which are well described in published guidelines.[1, 2] These include the specificity of training, such that competency in one procedure does not automatically confer competency in another and the notion that competency should be determined based on objective measures of skill rather than subjective impression of instructors or mere numbers of supervised cases performed by trainees. In fact, threshold numbers have come to signify minimum numbers of proctored examinations below which it is not possible to even assess a trainee for competency. Equally important is the principle that there ought to be one standard for what constitutes a competent examination that is specialty independent. For example, it is unethical to consider an individual with a cecal intubation rate of 75% competent when other trainees seeking to perform the procedure are consistently reaching the cecum in 85-90% of cases at the completion of their training.

Controversy over what constitutes sufficient training for a particular procedure and how many procedures to require trainees to perform with supervision can be resolved if there emerges the following two items:

a. A consensus as to what minimum level of performance on real procedures constitutes an acceptable definition of competence to perform the procedure independently in the community. Presumably this would derive from benchmarking data about clinical performance of the particular procedure by practicing endoscopists.

b. An assessment tool that could measure a trainees' skill and reliably predict whether the individual will be able to perform procedures at that minimal level of acceptable competency defined above.

Recently, investigators have validated such a tool for measuring trainee performance in colonoscopy on actual cases.[15] The development of a simulator-based assessment tool that could reliably predict competent performance would be of enormous value. It would allow an unbiased and reproducible measure and assure patients that those individuals performing their endoscopy, regardless of specialty, have been trained to sufficient standards of quality.

As part of the development of its Fundamentals of Endoscopic Surgery program, the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) has created global rating scales to measure clinical performance during gastroscopy and colonoscopy. These tools, referred to as Global Assessment of Gastrointestinal Endoscopic Skills (GAGES) were assessed in a multi-center trial including both surgeons and gastroenterologists.[16] The investigators were able to use GAGES to differentiate between novice and experienced operators, but the study was not powered to detect more subtle differences between different levels of training. Nor has GAGES yet been validated as being able to predict whether a trainee is able to perform independently complete competent colonoscopy to an agreed upon objective level of quality.

V. Methodology used for this PIVI

This PIVI was developed to propose thresholds for adopting simulators for training and assessing skill. This was done by first considering the current capabilities of simulators. Particular attention was given to the use of simulators for training in diagnostic colonoscopy and hemostasis of gastrointestinal bleeding, two areas of high clinical interest for which simulator applications have been investigated.

A comprehensive review of the medical literature of published trials using *ex vivo* and computer simulators was performed and circulated to the members of the PIVI committee. Study methodologies, size, and findings were considered by committee members for their relevance to the central questions and the level of evidence they constituted (Table 1). Given the readily-apparent limitations of sufficient literature pertaining to some of the key questions, this PIVI additionally relied upon expert opinion of its members. For that reason, the PIVI committee was

comprised of well-recognized investigators and thought leaders in the field of simulators and endoscopy education and credentialing. Participants included both gastroenterologists and surgical endoscopists with significant experience in the field of simulators and education.

Table 1. Levels of evidence

1a	Systematic reviews (meta-analysis) containing at least some trials of level 1b evidence, in which results of separate, independently conducted trials are consistent
1b	Randomized controlled trial of good quality and of adequate sample size (power calculation)
2a	Randomized trials of reasonable quality and/or of inadequate sample size
2b	Nonrandomized trials, comparative research (parallel cohort)
2c	Nonrandomized trial, comparative research (historical cohort, literature controls)
3	Nonrandomized, non-comparative trials, descriptive research
4	Expert opinions, including the opinion of PIVI committee members

VI. Literature review results/summary

a. Training

There is a substantial body of literature demonstrating the effectiveness of various endoscopy simulation devices for teaching endoscopic skills to trainees, especially in the early phases of training. Using the criteria for levels of evidence demonstrated in Table 1, there appears to be a good, but limited, number of scientific studies to suggest that there is a "degree" of benefit to the use of simulators in the instruction of flexible sigmoidoscopy, upper endoscopy, therapeutic colonoscopy, ERCP, and, perhaps, small bowel enteroscopy. Overall, it appears that simulators can be used to instruct trainees quickly and safely on the basic techniques, but do not replace live patient one-on-one mentorship. Despite the varied capabilities and promising potential, usage of the current endoscopic simulators appear to help primarily with early learning curves for endoscopic procedures, but have not yet translated into achievement of benchmarks for competency earlier in the training process or improved outcomes for patients.

The highest levels of evidence demonstrate benefit in the early phase of colonoscopy training without an ultimate shortening of the learning curve, as well as improvement in hemostasis skills following a series of intensive *ex vivo* hands on workshops.[4, 17,18] Individuals participating in three simulator hemostasis sessions performed better than those who did not attend these

workshops both on simulator skill assessment and on actual hemostasis rates on cases performed during the seven-month study period.[19] While corroborating studies will be important, the use of *ex vivo* models for hemostasis training appears to be the only area to date for which the literature supports current adoption of simulator work into the standard curriculum. To date, no papers have defined the standard learning curve for hemostasis skill of trainees or assessed the impact of *ex vivo* work on hastening competency in this area.

From the published experience to date, many other areas of simulator-based education appear to hold promise but will require more thorough investigation and validation to warrant calls for widespread adoption.

The current literature summarizing the usage of simulators for endoscopic training is delineated in Table 2:

Table 2. Literature on Simulators for Training

Author, year	Level of evidence	Subjects and simulator	Outcome (include type of final testing, clinical vs simulator)	Findings
Sigmoidoscopy				
Tuggy, 1998[20]	2a	Family Medicine Residents (10) Gastro-Sim (Interact Medical)	5 simulator trained versus 5 control Testing on 2 volunteers	-6-10 hours simulator significant improvement in time to splenic flexure, directional errors, overall exam time
Gerson, 2003[21]	1b	Internal Medicine Residents (16) Immersion Medical Simulator	9 simulator-mean 2 hours 7 bedside – 10 cases 5 test cases live patients	-Simulator group more difficulty with scope insertion, ability to reach flexure, retroflexion - No differences for satisfaction or times
Sedlack, 2004[22]	1b	Internal Medicine Residents (38) Immersion Medical Simulator	19 simulator 3 hrs + bedside 9 hours 19 bedside 12 hours 8-9 live cases	-Patient discomfort scores lower in simulator group - No differences in any performance parameters

Colonoscopy				
Sedlack, 2004[5]	1b	8 Novice GI Fellows Immersion Medical Simulator	4 Simulator training 6 hours 4 bedside training 15 live cases	-Simulator trained fellows outperformed bedside trained fellows in all aspects up to 30 cases except for time of insertion - 3 parameters showed advantage to 30 cases: depth insertion, independent completion, landmark identification
Cohen, 2006[4]	1b	45 First year GI fellows Simbionix GI Mentor	23 simulator for 10 hours 22 bedside 200 test cases	- Similar performance cases 1-20 - Simulator group better objective performance cases 21-80 - Simulator group better subjective performance cases 21-60 - No difference in discomfort scores
Park, 2007[23]	1b	24 General surgery and internal medicine residents Immersion Medical Simulator	12 Simulator for 2-3 hours 12 Background instruction only One live exam	Simulator group better global rating score on live examination
Snyder, 2010[24]	2b	13 surgical residents who had received training on Immersion Medical	Follow-up virtual colonoscopy case	Retention of skills on a virtual case 4.5 months of initial training

Summary of evidence on simulators for training

Flexible Sigmoidoscopy – No evidence of improved performance parameters comparing simulator training to bedside training. (Evidence – 1b). One trial suggesting improved patient satisfaction scores (Evidence – 1b). Comment: Time on Simulator only 2-3 hours, Immersion Medical.

Colonoscopy – Improved early learning curve (cases 21-80) but did not shorten the overall time to competency (Evidence-1b). Comment: Time on Simulator 10 hours, Simbionix. Another study demonstrated improved performance parameters on bedside cases up to 30 cases (Evidence-1b). Comment: Time on Simulator 6-10 hours, Immersion Medical.

Upper Endoscopy – Two clinical randomized control trials (RCTs) demonstrated superior performance on live cases either compared to no training or bedside training (Evidence-1b). Comment: Time on Simulator 5-10 hours, Simbionix.

Hemostasis - Improvement demonstrated for simulator-trained group (three sessions versus clinical training) but testing only on performance parameters on *ex vivo* simulator, no testing on bedside cases. Improved outcomes on actual hemostasis cases performed during the study period in the simulator trained group (Evidence-1b).

ERCP – One study demonstrated improved cannulation success and times on a mechanical simulator. No studies performed on bedside cases (Evidence-2b).

1. Simbionix found to be valid in one study (able to differentiate between expert and novice. Two ERCP cases performed by expert and novice based on mean total procedure time in one case and ability to cannulate the CBD plus procedure time for 2nd case.[25]

2. *Ex vivo* porcine simulator felt to be more realistic compared to Simbionix in comparison trial during a one- day therapeutic endoscopy course.[26]

3. Erlangen *ex vivo* porcine simulator found to be valid for ERCP simulation.[27]

EUS- No studies to date using an EUS simulator.

Laparoscopy- Practice on the Fundamentals of Laparoscopic Surgery simulator developed by SAGES was shown to improve performance in the operating room. Approximately 7.5 hours of simulator practice was needed for first year residents to perform at the level of third year residents in the operating room during laparoscopic cholecystectomy[11] (Evidence – A).

B. Assessing skill

There are only limited data examining the ability of endoscopy simulators to be used as tools to assess endoscopic skills. The limited literature available in this respect has shown these teaching tools to lack performance metrics of adequate sensitivity or reliability for use as meaningful

assessment tools. At best, some models have been able to distinguish novice from experts but generally have failed to differentiate gradations of skill level. To date, a simulator-based assessment tool with predictive validity has yet to be developed for any endoscopic procedure. The available literature examining simulation devices (computer and animal tissue models) in endoscopy skills assessment for some of the more common endoscopic procedures is summarized in Table 3. The level of evidence for each report is graded using the scale found in Table 1.

Colonoscopy

Owing to the routine training in colonoscopy by all GI fellows and general surgeon residents, this procedure has received the most attention by education researchers. For computer simulators, much of the assessment data is derived from studies performed over the past decade in attempt to provide initial validation of new commercially available devices. The earliest of these (Sedlack et al and Ferlitsch et al) examine the two early computer simulation models (Immersion Accutouch and Simbionix GI Mentor, respectively).[38, 39] In these two studies, a few of the computer generated metrics of a user's performance, primarily those related to time, appeared relevant, but the majority of the metrics did not appear to be significant measures of skill. This is primarily due to the relative ease of the simulation scenarios even for novice endoscopists as well as a lack of sufficient ability of the assessment metrics to detect a difference between groups of significantly different skill levels. Increasing the difficulty of the case has been shown to improve the accuracy and reliability of some performance metrics.[40] Multiple subsequent studies of these devices have yielded varied and mixed results, but all with essentially the same message; namely, that the metrics measured on currently available computer simulators cannot yet offer sufficient discrimination of operator skill to make meaningful predictions of competence on real procedures (Table 3). A more recent entry into the computer simulation market produced by Olympus has also produced mixed results from the assessment standpoint, [41,42] although it shows promise for both the case difficulty and the number of valid assessment metrics.

There are no data available on the use of live animal models for assessment in colonoscopy, although there are some data on the use of *ex vivo* animal models in this role. Although porcine organs are commonly used in centers with *ex vivo* model training, they are most often used for upper endoscopic procedures (EGD, ERCP, and EUS), as the porcine colon is quite short and thin-walled, limiting its usefulness for colonoscopic simulation. This limitation prompted the development of a bovine colon model. Based on published research, this bovine model can provide more difficult scenarios than computer simulation and, with the assessment scoring used, is able to differentiate skills based on a limited number of metrics such as cecal intubation rates, time to cecum, and percentage of mucosa visualized.[30] As part of ongoing research, this model is currently being used by the colorectal surgery specialty group at the Mayo Clinic (Rochester, Minnesota) to assess each of their advanced surgical fellows' colonoscopy skills at various stages of training.

Esophagogastroduodenoscopy

As with colonoscopy models, the relative ease of the available esophagogastroduodenoscopy (EGD) scenarios and insufficient sensitivity of the computer-measured performance metrics have precluded, to date, their usefulness as an assessment tool (Table 3). In a majority of studies, only novices could be distinguished from expert endoscopists, while the performance by intermediate skill groups were not measurably different from the other two experience levels.

The main advantage of *ex vivo* upper endoscopy models is the realism of tissue response to manipulation and therapeutic interventions that the haptics of computer simulation have yet to match. Despite extensive use of these models for training, there are still only limited data examining their usefulness in skill assessment. The only available study for use in assessment is a validation report of the Erlangen model by Neumann et al, [27] revealing that the use of this model in conjunction with a specific scorecard can differentiate novices from experts, similar to the computer models.

ERCP

The data on available ERCP computer simulators for assessment use is also quite limited. Bittner et al[25]performed a limited validation study of the GI Mentor II model demonstrating the simulator's ability to differentiate novice endoscopists from experts. Like the other types of computer simulation devices, though, this ability was limited to procedure times only, while the other computer performance parameters did not show measurable differences. Validation studies on two static models generated results similar to those of computer simulation where total procedure time appears to be the only measurable variable that distinguishes between experience groups.[43, 44] *Ex vivo* and live porcine models remain the main simulation platforms used in ERCP simulation training due to the need for realistic tissue pliability and response to cautery. To date, there are no published data regarding their use as assessment tools.

Endoscopic Ultrasound

Computer simulation in endosonography has only recently become commercially available and, as of yet, there are no data examining its use as an assessment tool. Nor are there data examining the potential of static, or porcine (*ex vivo* organs or live animal) models[45] for assessing skill.

Therapeutic Endoscopy

Although several of the computer models can simulate bleeding and polypectomy, there is no evidence for their use in either training or assessment of these techniques. *Ex vivo* models have been preferred for training due to similar requirements for ERCP such as tissue pliability and response to therapy. The majority of the data exists for assessment of hemostasis techniques (injection, clip, diathermy, band ligation). There are several well-conducted studies of reasonable power that have used checklists to establish training efficacy. However, there have been no

published studies investigating the validity of these checklists. One study has investigated other therapeutic techniques (polypectomy, stricture dilatation, PEG insertion), but again used unvalidated checklists.[46]

Table 3. Validation Studies for EGD, Colonoscopy and Sigmoidoscopy

Author, year	Level of evidence	Subjects and Simulator	Outcome (include type of final testing, clinical vs simulator)	Findings
Phitayakorn (2008)[28]	3	GI Mentor II 23 expert surgeons (>1000)	Time to cecum, time to completion, % mucosa visualized, % time looped.	Wide range of variability for all measurements.
Koch (2008)[29]	2b	Olympus Endo TS-1 23 experts (>1000), 26 novices (0)	Dexterity score, time to cecum, max shaft insertion force, shaft torque, tip section force, max pain.	Time to cecum only discriminated.
			Expert judgment of face validity.	All scores 6.9/10 or higher.
Sedlack (2007)[30,31]	2b	Bovine *ex vivo* 13 experts (?), 13 intermediate (100-150), 13 novices (0)	Unvalidated blueprint - time to markers.	Completion rate, completion time, time to markers, % mucosa visualized, quality of mucosa exam all discriminate.
			Expert judgment of face validity.	All except one (haustra realism) better than neutral
Koch, 2007 (29)	2b	GI Mentor II 35 experts (>1000), 20 experienced (200-1000), 15 intermediate (<200), 35 novices (0)	Dexterity score, time to cecum, % time with clear view, lost view of lumen, excessive local pressure, % time in pain, excessive loop formed.	Novices distinguished on dexterity task, completion time, visibility and patient pain. No significant differences found between

				intermediate, experienced, and experts on any parameters.
Westman (2007)[32]	3	GI Mentor II 11 experts	Visuospatial tests (PicSOr) Expert opinion.	The endoscopists who performed better in the visuospatial tests also were better at maintaining visualization of the colon lumen. Those who performed better in the PicSOr test formed fewer loops during colonoscopy.
Grantcharov (2005)[33]	2b	GI Mentor II 8 experienced (>200), 10 residents (<50), 10 novices (0)	Time to cecum, % time with clear view, lost view of lumen, excessive local pressure, % time in pain, excessive loop formed.	Novices distinguished on completion time, % mucosa seen, efficiency of screening, patient pain, loop formation, time with loop, excessive local pressure. No sig differences found between experienced and experts.
Haycock (2009)[34]	2b	Olympus Endo TS-1 11 experts (>1000), 13 intermediate (<1000), 10 novices (0)	40 parameters divided into general measures, technical measures, patient aspects, time scores, looping measures.	Novices distinguished between trainees and experts on 18 measures. Trainees distinguished between experts on 5 measures (all time scores).

Eversbusch (2004)[35]	2b	GI Mentor 8 experts (>200), 10 residents (<50), 10 novices (0)	Time to cecum, % mucosa visualized, % time with clear view, efficiency of screening, excessive local pressure, pain, time in pain, loop formation, time with loop.	Learning curve plateau on 2nd repetition for experts, 5th for residents, 7th for novices.
Sedlack (2003)[36]	2b	Accutouch 10 experts (?), 6 fellows (150), 6 residents (0)	Total time, insertion time, withdrawal time, time in red-out, max insertion depth, total path length, % mucosa visualized, % discomfort, volume air insufflated, max scope force, complications, extra sedation, identification of pathology.	Time, insertion time and red-out distinguished novices from residents. No significant differences found between residents and experts.
			Expert judgment of face validity.	All better than neutral.
Mahmood (2003)[37]	2b	HT Immersion (Accutouch) 5 experts (>100), 7 intermediate (11-100), 10 novices (<10)	Total time, % mucosa seen, path length, perforation.	Total time, % mucosa seen, perforation discriminated overall, but no direct comparisons made between the three groups.
Ferlitsch (2002)[38]	2b	GI Mentor 11 experts (>1000), 13 novices (0)	Endobasket, endobubble scores, correctly identified pathology, insertion time, successful retroflexion, adverse events, excessive wall pressure, impaired view.	Experts significantly better than novices.

Sedlack (2002)[39]	2b	Accutouch 10 experts (?), 5 partially trained (?), 2 novices (0)	Time to complete, distance intubated, % mucosa seen, complications, pain.	Average and minimal performance standards established. Simulator likely most effective early in training.

Appendix 1. Study designs needed to address the questions raised by this PIVI.

In the evaluation of simulators, it is expected that study methodology will address both technical and cognitive aspects of the procedure that is being simulated. The methods for evaluating simulators will depend on the goals of the simulator (training vs. assessment of competency).

Training

Simulators developed primarily for training purposes should be relatively low in cost such that they can be widely available. It is unlikely that simulators developed primarily for training will replace clinical training and experience. The goal of training simulators should be to decrease the number of clinical procedures needed to reach clinical proficiency, with the understanding that a certain amount of clinical training and experience will always be necessary for any procedure. Therefore, the relevant outcome is whether a given training simulator reduces the number of actual clinical procedures needed to reach clinical proficiency for a given endoscopic procedure. The optimal endpoint of training simulators should be number of clinical procedures needed to assess for competency. For some procedures such as colonoscopy, professional societies recommend specific minimal threshold numbers of cases which must be performed under supervision before competency reasonably can be assessed, based on published prospective data. However, for many other procedures where the evidence supporting current recommended thresholds is less robust, such as the application of clips to control ulcer bleeding, validation of the value of simulators will first need to have established the standard learning curve for the particular technique. Provided this information is known, the benefit of using simulators for training can only be validated in an adequately powered controlled trial that demonstrates that the incorporation of simulators in a training program leads to acquisition of technical and cognitive competency with fewer clinical endoscopic procedures than required using supervised endoscopic instruction without access to the simulator.

While shortening the learning curve for trainees to reach objective levels of proficiency will be the primary endpoint for future validation study, another potential endpoint that would promote adoption of simulators into standard curricula would be demonstration of significant cost and manpower reductions for instructors during the training process. Future studies will need to

consider the transfer efficiency ratio to determine if modest benefits observed justify the particular time and costs involved.

Future studies to validate a role for simulators in training will also need to address issues of generalizability. Clearly the potential benefits of simulator-based work may depend as much on the particular baseline experience level of the trainee as the attributes of the educational exercise itself.

Assessment

A simulator that is able to assess the actual clinical skills (technical and cognitive) of a clinician performing endoscopy may be a useful tool for credentialing (or recredentialing). A prerequisite for any validation of such a tool would be a prior consensus of how to measure competency for a particular procedure in real patients.

Simulators designed for assessing clinical skills can only be valid if performance on the simulator is correlated with existing accepted clinical performance metrics on actual patients. For example, a study that evaluates the ability of a colonoscopy simulator to assess the clinical skills of an endoscopist should determine whether performance on the simulator is correlated with an endoscopist's cecal intubation rate or adenoma detection rate as endpoints in an adequately powered and controlled study. Any study that evaluates competency needs to include endoscopists at different levels of training and not just constrain itself to studying endoscopists at the two ends of the competence spectrum. At a minimum, a simulator designed for assessing endoscopic skill must be able to distinguish endoscopists with beginning level (equivalent to a first year trainee), intermediate level (average second year trainee), fully trained (third year trainee or recent graduate), and expert level (endoscopic instructor with at least five years training experience) endoscopic skills. Furthermore, the evaluation should be correlated with actual clinical outcomes to validate the performance of the simulator. In the opinion of this committee, it is not acceptable to rely upon the results of a high-stakes examination on a simulator in which the results are measured only against how well experts perform on the simulator itself.

The use of simulator-based assessment as an educational tool to guide the learning process ought to be validated similarly to simulator-based training in protocols to test the impact of such efforts upon the learning curve to acquisition of established outcomes, as outlined above.

Appendix 2. Technological advances needed to achieve the thresholds for this PIVI.

In general terms, to develop simulators that can achieve a 25% or more reduction in the learning curve for particular endoscopic procedures, investigators and industry will need to focus on what skill components are learned at what stage in the development of competency. It may turn out that deconstruction of techniques into specific skill sets will lead to faster learning progression. Rather than one ideal simulator, the optimal path to competency may turn out to be a number of

different tools to help trainees develop a variety of skills that are then integrated on actual proctored procedures. Several technological advances may allow simulators to achieve the goals of this PIVI. These would include improved realism, enhanced ability to modify exercises to increase difficulty with repeated practice and increased integration of pathological recognition to hasten the acquisition of the cognitive aspects of competency. However, just as important will be efforts to better deconstruct various endoscopic procedures into the component steps and requisite skills. Progress in this area may be necessary to guide simulator development sufficiently to achieve the thresholds of this PIVI.

What is needed for a simulator based skills assessment to meet the threshold of this PIVI for its use in credentialing and recredentialing?

The utility of a device as an assessment tool is based on a number of factors, any one of which can result in poor accuracy and reliability as an assessment tool. The first factor is establishing an appropriate level of complexity for the simulation scenario to match the group being assessed. Cases that are too easy, as many of the computer simulation scenarios have been shown to be, may be ideal for helping novices train toward competency though may not lend itself to being a sensitive tool for assessment. The second major factor that may limit a simulator's usefulness hinges not on the simulation itself but the metrics and methods used for grading. For computer simulation, this involves the programmed formulas that attempt to assign a performance score to a given metric. Lack of meaningful parameters or appropriate calculation of scores will lead to results of little utility. Similarly, for animal models, this relates to the scoring form and parameters used. As suggested by the current body of literature, parameters related to time are relatively straightforward and objective and, as a result, the most useful regardless of the device used. Other parameters, such as scope force or percent mucosa visualized, rely on more complex formulas or more subjective scorer assessment and have proven less sensitive metrics though remain desirable goals.

Another significant factor in developing an adequate set of metrics for these more subjective skill parameters has been a lack of defined performance expectations for the spectrum of user's experience. As educators, we naïvely reassure ourselves that we know competency endpoints when we see them, but this is not objective nor does it allow for establishing expected learning curves for specific skills. As such, we may have difficulty establishing when a trainee is ahead of the learning curve or falling behind except when the variance becomes egregious. Many questions need to be answered for each individual assessment, for example: "How often should a novice, intermediate, or expert be able to complete a specific task and how long should it take them?" or "What percent of mucosa does each group typically view during an exam?" Without the establishment of an objective set of desired performance endpoints by the GI community for various levels of experience, programmers or model developers have little information to rely upon to aid them in setting an appropriate level of difficulty for simulation models or finding the appropriate sensitivity for assessment parameters.

Within colonoscopy, there have been some recent developments in methods for fellow skills assessment during patient-based procedures and the identification of learning curves.[15, 47,48] Specifically, use of the Mayo Colonoscopy Skills Assessment Tool (MCSAT) has demonstrated the learning curves for many of the core motor and cognitive skills required for colonoscopy and has established minimal competency thresholds for these metrics.[15] Hopefully these data, or other similar information, may lead to more precise learning curve expectations for various skills and act as a guide by which model developers can refine current or future models and assessment. If so, similar data would be needed in other endoscopic procedures to eventually reach the ultimate goal of having simulation assessment metrics that correlate well with patient-based skill assessment.

Summary of suggestions for future development and evaluation:

1. Development of accurate, reliable and validated clinical patient-based assessment tools for all endoscopic procedures to benchmark clinical performance abilities and learning curves.

2. Development of simulated cases of increased difficulty to allow discrimination of differing grades of ability (not just novice/expert)

3. Development of clinically relevant, accurate and reliable simulator assessment metrics that correlate with clinical performance benchmarks

4. Simulator assessment metrics to be validated to establish efficacy in discriminating differing grades of ability (not just novice/expert). Establishment of the following for each assessment task/module:

    a. Learning curve

    b. Benchmarks of performance (using expert performance)

Specific future needs in the field of endoscopy simulation:

> *What factors should guide investigators and industry as they aim to develop simulators that will meet the thresholds of this PIVI?*
>
> *Which particular procedures and which particular trainees should be the focus?*
>
> *What technological advances are most needed to produce models that will allow thresholds to be met?*

As professional societies work to better define the constituents and benchmarks for competency in various procedures, investigators and providers of simulators will need to focus their efforts on addressing current unmet needs, refining which trainees get the most benefit from which tools, and reducing the cost of simulation to improve access and use.

What are the biggest unmet needs in the area of endoscopic training today?

To some extent this may depend upon geography, local manpower issues for endoscopic trainers, and certainly upon the case volume for a procedure in question. For colonoscopy, in most regions where practitioners with cecal intubation rates well over 90% are readily available, the need is to develop ways to ensure trainees going into independent practice can achieve benchmark rates of adenoma detection and removal. There is increasing demand for finding opportunities for practicing endoscopists to learn new techniques and refresh existing skills. There is scant literature to date on the use of various simulators for maintaining skill in different procedures; while hands-on *ex vivo* workshops have proliferated in recent years for post-graduate attendees, no systematic information has been collected to validate and optimize the benefit from such effort. Finally, as many procedures move further away from the purely diagnostic and as therapeutic intervention is increasingly integrated into procedures, greater attention may be needed to maximize the training and retraining in therapeutic skills and the cost effective use of simulators to this end.

Ex vivo models currently are capable for instruction in most therapeutic techniques. Limitations remain the inability to simulate bowel motility, the simulation of submucosal blood flow, and bleeding during resection training. For many procedures that can be simulated on animal tissue models; however, there is a relative dearth of material delineating how best to teach the necessary skills using the model, including detailed breakdown of techniques into components for demonstration, practice, and assessment with feedback. Train-the-trainer educational materials and validated programs will be advantageous.

For computer simulators, there is consensus that the models need to become more realistic and maneuvers on them made more difficult to perform. This is particularly true for therapeutic techniques, which currently lag far behind efforts with *ex vivo* models. Haptic improvement in the "feel" of accessories will be required as will the simulation of non-linear changes in resistance to scope passage, for example as loops are created and reduced. There is further potential for increased use of visible human 3-D anatomy outside the bowel and integration into training modules of virtual cases of increasing complexity and anatomical variety. Perhaps the biggest untapped potential of the computer simulators available to date is the ability to integrate interactive quizzing with immediate feedback on matters of lesion recognition, appropriate management decisions based on findings, and specific selection of accessories and generator settings to perform therapeutic maneuvers.

There is no doubt that technological advances will expand the potential realism of the simulator experience, but the cost and efficacy questions currently give pause to consider whether this is desirable or necessary. With much of the evidence to date (see above) suggesting only a modest improvement in the early phase of learning for computer simulators (in EGD and colonoscopy), for those procedures with ample real patient training opportunities available, a more compelling interest would be for innovation to produce simple models without increased realism, but at a

much lower cost. Important technological innovation is needed in the area of finding a way to surmount the cost obstacle to wider access to simulator-based training. This could be achieved via the development of improved low-end static models, by very basic virtual reality simulators geared for novice trainees, or by the development of training centers at regional sites to allow students access to otherwise cost prohibitive high-end realistic computer simulators.

There has been far too little emphasis on cognitive training using simulators rather than the technical aspects of hand-eye coordination, getting to the cecum, and pulling out loops. To that end, a major effort is needed to develop and maintain a large database of cases for web-based interactive learning with stop-and-go video interlaced with interactive quizzing. Such virtual instructors could fulfill the unmet needs of cognitive training. The integration of enhanced tools that would facilitate students' progress beyond the cognitively incompetent albeit technically proficient level should be a chief focus of technological development with respect to computer-based endoscopy simulation.[49]

Appendix 3. Cost of Current Endoscopy Simulation

There are many forms of simulators from animal (*in vivo* and *ex vivo*) to computer-based mechanical, to models and the cost a few thousand dollars for the simplest of models up to and exceeding $200,000 for complex multipurpose simulators (Table 4). At the current level of development, most simulators require only a brief period of training followed by live patient training. Given current costs of simulators and available data presented above regarding improvement in clinical outcomes, usage of computer simulators does not currently appear to be cost-effective for training in endoscopy.

Proponents of simulators point to the potential cost savings by virtue of freeing up resources and time of instructors; however, this yet unrealized potential assumes that simulator work will facilitate independent instruction of trainees. Whether or not maximal benefit of simulators will require instructor input has not yet been determined.

It is likely the development and validation of reliable simulator-based assessment tools will increase the demand for simulators and this may greatly impact the amount of resources available to offset their cost.

Table 4: Cost of Available Simulators

Manufacturer	Name	Type	Modules	Price	Comments
Medical Innovations	Endo X Trainer	Composite Plastic/ Animal	EGD/ Colonoscopy/ ERCP/ Polyps/	$3449	Erlangen-type model with porcine organs

			Bleeding Porcine Organ Package		
International Inc.					
Endosim, LLC	EASIE-R	Composite Plastic/ Animal	EGD / Colonoscopy/ ERCP/Polyps	$2100	Erlangen-type model with porcine organs
			EUS/Roux-en-Y ERCP/ Enteroscopy Porcine Organ Package		Organ packages in various configurations must be purchased separately; available for $250 - 750.
Delegge Medical	Various Erlangen-Type Endoscopy Models ("Endo Billy," "Endo Eddy," etc.)	Composite Plastic/Animal	EGD / Colonoscopy/ ERCP/Polyps	$2250-2850	Erlangen-type model with porcine organs
			GI bleeding		Organ packages in various configurations must be purchased
			Porcine Organ Package		Separately; available for $135 - 250.
Simbionix	GI Bronch Mentor	Virtual-Reality	Gastroscopy & Colonoscopy modules	$64,500	Standard package, does not include duodenoscope
			Gastric Emergency Bleeding Module	$6000	
			Flexible Sigmoidoscopy	$5000	
			Duodenoscope	$6000	Required for ERCP and EUS Modules
			ERCP Modules (two)	$15,000	Includes both modules
			EUS Module	$15,000	
			Shipping and handling	$2500	Installation, training, calibration, 1-yr. warranty included
			Total	$114,000	
CAE Healthcare	AccuTouch	Virtual-Reality	UGI package	$76,750	Includes UGI scope head, EGD module, and ERCP module

	LGI package	$45,750	Added to UGI package, includes LGI scope head, flex sig module, flex sig supplemental module, colonoscopy module, colonoscopy biopsy module, and polypectomy module. Includes 2 22-inch, flat-panel monitors, computer
	Upper GI bleed upgrade package	$19,000	
	Total	$141,500	
Modules can be purchased separately:	Intro to Flex Sig Module	$7175	
	Flex Sig Supplemental Module	$7175	
	Colonoscopy Intro Module	$7175	
	Colonoscopy Biopsy Module	$8650	
	Colonoscopy Polypectomy Module	$8650	
	Intro to EGD Module	$8650	
	ERCP Module	$8650	

Appendix 4. Methods of simulator validation

One of the limitations in proving the value of simulation has been the lack of a defined method of validation. To date, many simulator studies have taken an "off-the-shelf" simulator that has been designed with the intent of "looking" real and tested it in a variety of ways to attempt to prove the simulator has value. If the following template were used for simulator development and validation, better simulators would result and more reliable results could be obtained.

Task deconstruction. The first step of designing a simulator is to clearly define the task that is being simulated by deconstructing it into basic skills. For instance, the motor aspects of flexible GI endoscopy can be deconstructed into five skill sets: 1) Navigation, 2) loop reduction, 3) retroflexion, 4) mucosal evaluation, 5) targeting. This deconstructed task list is developed by a consensus of experts. Each task should be as unique as possible so that performance in one domain can be separated from another.

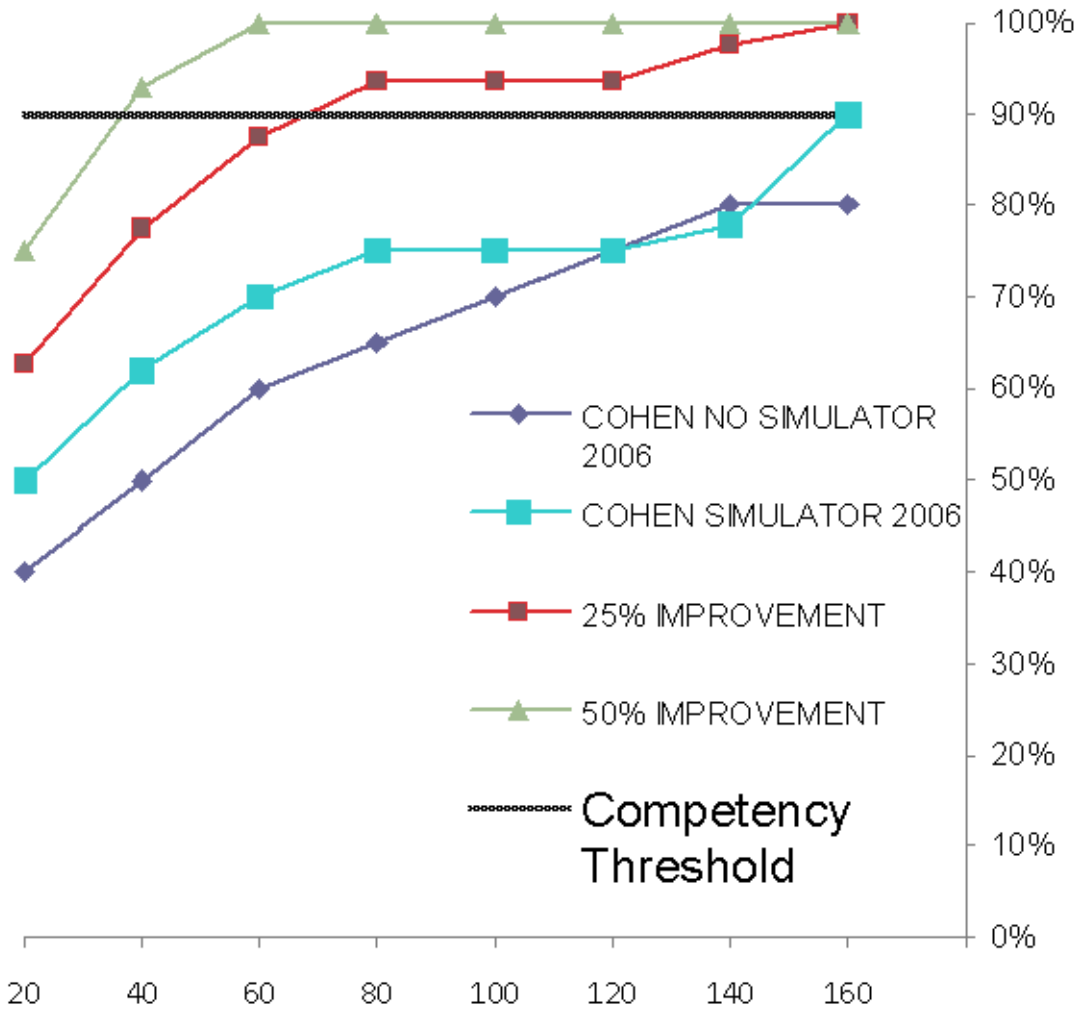
Metrics. After developing the deconstructed task list, it must be decided on how performance in each domain will be measured. Metrics must take into account important factors for proficiency and not simply depend on time.

Simulator design. Once the deconstructed tasks and metrics have been defined, the simulator is then designed. It is not necessary to make the simulator look like the "real" clinical environment and, in fact, it is often advantageous to work in a "game" environment without the distractions of how well the "real" environment has been created. Each simulated task should be as unique as possible from the next so that performance in one domain impacts performance in another as little as possible. The metrics of performance measurement must be taken into account during the design process as well to assure that accurate measurement will be possible. Note that this process of design is in reverse of that typically used in evaluating "off-the-shelf" simulation platforms designed to look "real" but without the background of a deconstructed task list and well-defined metrics.

Validation. After designing a simulator, its utility in both transferring and assessing the intended skills must be validated. There are five primary types of validation – face, content, construct, concurrent and predictive. Each of these validated parameters should be assessed in multi-institutional studies with participation of all relevant specialties (ie, surgery and medicine for flexible GI endoscopy).

One of the most important goals of simulation is to achieve predictive validity, ie, does performance on the simulator predict performance in the real clinical environment. In order to establish this level of validity, it is important that a validated clinical assessment tool be used. All currently published studies on flexible GI endoscopy have compared performance on the simulator to clinical performance without the benefit of a validated clinical assessment tool. This may be one of the most important factors in why it has been difficult to prove the value of simulation to date. The literature suggests that a global assessment of performance as judged by an expert evaluator is a more reliable assessment tool than one that relies on more specific measurements such as time, percent mucosa inspected, etc. A few published endoscopy skills assessment tools exist with varied degrees of validity evidence. The Mayo Colonoscopy Skills Assessment Tool (MCSAT), Direct Observation of Procedural Skills scores (DOPS), and Global Assessment of Gastrointestinal Endoscopic Skills (GAGES) each measure various aspects of deconstructed colonoscopist skills.[16, 50,51]

Figure 1. Effects on Learning Curves for Colonoscopy Skill Acquisition Using Simulators of Increasing Efficacy



The first two curves are adapted from a randomized multicenter trial of GI Mentor II vs. no simulator experience. [2].

VII. PIVI Committee

Jonathan Cohen, MD, Committee Chair, Lauren B. Gerson, MD, Brian Bosworth, MD, Amitabh Chak, MD, Brian Dunkin, MD, Dayna Early, MD, Robert Hawes, MD, Adam Haycock, MD, Juergen Hochberger, MD, Joo Ha Hwang, MD, John Martin, MD, Peter McNally, MD, Robert Sedlack, MD, and Melina Vassiliou, MD

Special acknowledgement to Brooks Cash, MD, and Betsy Rodriguez, MD, for the review and support of this document.

VIII. REFERENCES

1. Wexner SD, Litwin D, Cohen J, et al. Principles of privileging and credentialing for endoscopy and colonoscopy. Gastrointest Endosc 2002;55:145-8.
2. Wexner SD, Eisen GM, Simmang C. Principles of privileging and credentialing for endoscopy and colonoscopy. Surg Endosc 2002;16:367-9.
3. Cohen J, Greenwald D. Training in endoscopy: a historical perspective. In Successful Training in Endoscopy. J Cohen, Editor. Wiley-Blackwell. 2011:3-15.
4. 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;64:361-8.
5. Sedlack RE, Kolars JC. Computer simulator training enhances the competency of gastroenterology fellows at colonoscopy: results of a pilot study. Am J Gastroenterol 2004;99:33-7.
6. Fraenkel J, Wallen N. Correlation research. In: How to design and evaluate research in education. 5th ed. 336-63. McGraw Hill, New York. 2003.
7. Beyer L, Troyer JD, Mancini J, et al. Impact of laparoscopy simulator training on the technical skills of future surgeons in the operating room: a prospective study. Am J Surg 2011; 202:265-72.
8. Santos BF, Reif TJ, Soper NJ, et al. Effect of training and instrument type on performance in single-incision laparoscopy: Results of a randomized comparison using a surgical simulator. Surg Endosc. 2011;25:3798-804
9. Vassiliou MC, Feldman LS, Andrew CG, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. Am J Surg 2005;190:107-13.
10. Fried GM, Feldman LS, Vassiliou MC, et al. Proving the value of simulation in laparoscopic surgery. Ann Surg 2004;240:518-25; discussion 525-8.
11. Sroka G, Feldman LS, Vassiliou MC, et al. Fundamentals of laparoscopic surgery simulator training to proficiency improves laparoscopic performance in the operating room-a randomized controlled trial. Am J Surg 2010;199:115-20.
12. Gerson LB. Evidence-based assessment of endoscopic simulators for training. Gastrointest Endosc Clin N Am 2006;16:489-509, vii-viii.
13. Nelson DB, Bosco JJ, Curtis WD, et al. ASGE technology evaluation report. Endoscopy simulators. May 1999. American Society for Gastrointestinal Endoscopy. Gastrointest Endosc 1999;50:935-7.
14. Inaba K, Recinos G, Teixeira PG, et al. Complications and death at the start of the new academic year: is there a July phenomenon? J Trauma 2010; 68:19-22.
15. Sedlack RE. Training to competency in colonoscopy: assessing and defining competency standards. Gastrointest Endosc 2011; 74:355-66.
16. 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 2010;24:1834-41.
17. 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 with clinical endoscopy training alone. Gastrointest Endosc 2005;61:204-15.

18. 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;70:835-45.
19. 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: a prospective trial. Eur J Gastroenterol Hepatol 2006;18:1217-25.
20. Tuggy ML. Virtual reality flexible sigmoidoscopy simulator training: impact on resident performance. J Am Board Fam Pract 1998;11:426-33.
21. Gerson LB, Van Dam J. A prospective randomized trial comparing a virtual reality simulator to bedside teaching for training in sigmoidoscopy. Endoscopy 2003;35:569-75.
22. Sedlack RE, Kolars JC, Alexander JA. Computer simulation training enhances patient comfort during endoscopy. Clin Gastroenterol Hepatol 2004;2:348-52.
23. 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.
24. Snyder CW, Vandromme MJ, Tyra SL, et al. Retention of colonoscopy skills after virtual reality simulator training by independent and proctored methods. Am Surg;76:743-6.
25. Bittner JG, Mellinger JD, Imam T, et al. Face and construct validity of a computer-based virtual reality simulator for ERCP. Gastrointest Endosc 2010;71:357-64.
26. Sedlack R, Petersen B, Binmoeller K, et al. A direct comparison of ERCP teaching models. Gastrointest Endosc 2003;57:886-90.
27. Neumann M, Mayer G, Ell C, et al. The Erlangen Endo-Trainer: life-like simulation for diagnostic and interventional endoscopic retrograde cholangiography. Endoscopy 2000;32:906-10.
28. Phitayakorn R, Marks JM, Reynolds HL, et al. Expert benchmark for the GI Mentor II. Surg Endosc 2009;23:611-4.
29. 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;22:158-62.
30. Sedlack RE, Baron TH, Downing SM, et al. Validation of a colonoscopy simulation model for skills assessment. Am J Gastroenterol 2007;102:64-74.
31. Sedlack RE. Validation of computer simulation training for esophagogastroduodenoscopy: Pilot study. J Gastroenterol Hepatol 2007;22:1214-9.
32. Westman B, Ritter EM, Kjellin A, et al. Visuospatial abilities correlate with performance of senior endoscopy specialist in simulated colonoscopy. J Gastrointest Surg 2006;10:593-9.
33. Grantcharov TP, Carstensen L, Schulze S. Objective assessment of gastrointestinal endoscopy skills using a virtual reality simulator. Jsls 2005;9:130-3.
34. Haycock AV, Bassett P, Bladen J, et al. Validation of the second-generation Olympus colonoscopy simulator for skills assessment. Endoscopy 2009;41:952-8.
35. 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;18:1514-8.
36. Sedlack RE, Kolars JC. Validation of a computer-based colonoscopy simulator. Gastrointest Endosc 2003;57:214-8.
37. Mahmood T, Darzi A. A study to validate the colonoscopy simulator. Surg Endosc 2003;17:1583-9.

38. Ferlitsch A, Glauninger P, Gupper A, et al. Evaluation of a virtual endoscopy simulator for training in gastrointestinal endoscopy. Endoscopy 2002;34:698-702.
39. Sedlack RE, Kolars JC. Colonoscopy curriculum development and performance-based assessment criteria on a computer-based endoscopy simulator. Acad Med 2002;77:750-1.
40. Fayez R, Feldman LS, Kaneva P, et al. Testing the construct validity of the Simbionix GI Mentor II virtual reality colonoscopy simulator metrics: module matters. Surg Endosc 2010;24:1060-5.
41. Haycock AV, Bassett P, Bladen J, et al. Validation of the second-generation Olympus colonoscopy simulator for skills assessment. Endoscopy 2009;41:952-8.
42. Koch AD, Haringsma J, Schoon EJ, et al. A second-generation virtual reality simulator for colonoscopy: validation and initial experience. Endoscopy 2008;40:735-8.
43. Leung JW, Lee JG, Rojany M, et al. Development of a novel ERCP mechanical simulator. Gastrointest Endosc 2007;65:1056-62.
44. 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;7:389-96.
45. Matsuda K, Hawes RH, Sahai AV, et al. The role of simulators, models, phantoms. Where's the evidence? Endoscopy 2006;38(Suppl 1):s61-4.
46. Haycock AV, Youd P, Bassett P, Saunders BP, Tekkis P, Thomas-Gibson S. Simulator training improves practical skills in therapeutic GI endoscopy: results from a randomized, blinded, controlled study. Gastrointest Endosc. 2009;70:835-45.
47. 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, 1133 e1-3.
48. Thomas-Gibson S, Bassett P, Suzuki N, et al. Intensive training over 5 days improves colonoscopy skills long-term. Endoscopy 2007;39:818-24.
49 . Neumann M, Siebert T, Rausch J, et al. Scorecard endoscopy: a pilot study to assess basic skills in trainees for upper gastrointestinal endoscopy. Langenbecks Arch Surg 2003;387:386-91
50. Sedlack RE. The Mayo Colonoscopy Skills Assessment Tool: validation of a unique instrument to assess colonoscopy skills in trainees. Gastrointest Endosc 2010;72:1125-33,1133 e1-3.
51. Thomas-Gibson S, Bassett P, Suzuki N, et al. Intensive training over 5 days improves colonoscopy skills long-term. Endoscopy 2007;39:818-24.