Identifying and reducing errors with surgical simulation

M P Fried, R Satava, S Weghorst, A G Gallagher, C Sasaki, D Ross, M Sinanan, J I Uribe, M Zeltsan, H Arora, H Cuellar

Qual Saf Health Care 2004;13(Suppl 1):i19-i26. doi: 10.1136/qshc.2004.009969

The major determinant of a patient's safety and outcome is the skill and judgment of the surgeon. While knowledge base and decision processing are evaluated during residency, technical skills—which are at the core of the profession—are not evaluated. Innovative state of the art simulation devices that train both surgical tasks and skills, without risk to patients, should allow for the detection and analysis of errors and “near misses”. Studies have validated the use of a sophisticated endoscopic sinus surgery simulator (ES3) for training residents on a procedural basis. Assessments are proceeding as to whether the integration of a comprehensive ES3 training programme into the residency curriculum will have long term effects on surgical performance and patient outcomes. Using various otolaryngology residencies, subjects are exposed to mentored training on the ES3 as well as to minimally invasive trainers such as the MIST-VR. Technical errors are identified and quantified on the simulator and intraoperatively. Through a web based database, individual performance can be compared against a national standard. An upgraded version of the ES3 will be developed which will support patient specific anatomical models. This advance will allow study of the effects of simulated rehearsal of patient specific procedures (mission rehearsal) on patient outcomes and surgical errors during the actual procedure. The information gained from these studies will help usher in the next generation of surgical simulators that are anticipated to have significant impact on patient safety.

tasks. All these concerns have a direct impact on the quality and safety of patient care.

At over 300 000 operations per year, endoscopic sinus surgery (ESS) is one of the most common procedures undertaken by otolaryngologists in the USA. It also carries significant risk to the surrounding anatomical structures, such as the eye and the brain. Overall rates for major and minor complications as a result of ESS vary, but most studies have reported a rate of between 5% and 10%.^{1 2} Malpractice suits have awarded large sums to patients for loss of vision or the sequelae of iatrogenic cerebrospinal fluid leakage.

Although the concept of ESS is quite straightforward, performing the procedure skillfully and safely can be challenging.³ The relevant anatomy is complex and compact, with the added concern of having very important structures such as the brain, orbital contents, and carotid artery closely juxtaposed and therefore at risk.^{2 4} Thus there is very little room for error. In addition, the surgeon must navigate and manipulate this environment with both dominant and non-dominant hands, while coordinating their movements indirectly with the aid of a television monitor. Well developed hand-eye coordination is an obvious prerequisite.

Currently, the training curriculum of otolaryngology residents in ESS includes videotapes, cadaver dissection (where available), and direct observation of procedures in the operating room. As residents progress in their training, they are given a more active role in the procedure, ultimately becoming the major participant by their final year.

Although simulation training has been a core technology for aviation safety, new virtual reality (VR) simulations are an innovative approach to surgical training, one which will revolutionise education and error reporting in the healthcare field. Applying the technology and methods that have proven effective in aircraft pilot training may significantly improve surgical procedure training. Over the past three decades, computer assisted devices have had significant success in augmenting the education and training of surgical residents in several fields.⁵⁻¹⁰ VR simulation has in fact already played an introductory role in the training of residents for laparoscopic, gastrointestinal, plastic, ophthalmological, dermatological, and some laryngological procedures.^{9 11-16} The efficacy of VR simulation as a teaching tool is clear, but whether it is superior

Patient care requires multiple diagnostic and therapeutic endeavours, often invasive. Non-invasive laboratory tests can be evaluated for safety, reproducibility, and efficacy, but technical procedures are rarely measured objectively. Indeed, surgical training programmes (residencies) rarely, if ever, evaluate medical students' manual dexterity as a criterion for admission. The resident's innate ability is not assessed at the outset of training, and it is hoped that it is acquired during the years of the programme. The issue of measuring progress is critical and one that teachers of surgery have not resolved. Moreover, the acquisition of skills is seldom done with inanimate objects or devices, requiring live patient experience to master the

See end of article for authors' affiliations

Correspondence to:
Dr Marvin P Fried,
Montefiore Medical
Center, Department of
Otolaryngology, 3400
Bainbridge Ave 3rd Floor,
Bronx, NY 10467;
mfried@montefiore.org

Abbreviations: AHRQ, Agency for Healthcare Research and Quality; ES3, endoscopic sinus surgery simulator; ESS, endoscopic sinus surgery; MIST-VR, minimally invasive surgical trainer virtual reality; PicSOR, pictorial surface orientation; VR, virtual reality

to conventional teaching methods remains uncertain in many instances.^{11 12 15}

As virtual reality surgical simulators are developed, they must be shown to be both instructionally effective—able to teach the real skills needed for surgery—and valid in their evaluation of surgical skill. Although many simulators have been developed, only a few have been formally studied for their effectiveness in training and evaluation. Most of these validation efforts have compared experienced surgeons to those with less or no surgical experience in order to demonstrate construct validity—showing that the simulator is measuring the skill it is designed to measure. The results are mixed, with an anastomosis simulator,⁸ a knee arthroscopy simulator,¹⁷ the endoscopic sinus surgery simulator,¹⁸ and a laparoscopic skill simulator¹⁹ showing significant differences among subjects with different degrees of surgical experience, while other simulators have failed to discriminate among those with different levels of experience.^{20 21}

So far, there has only been a limited number of studies looking beyond construct validity in their evaluation of surgical simulators. The minimally invasive surgical trainer virtual reality (MIST-VR) laparoscopic simulator has been shown to possess construct validity, as experienced surgeons perform better than novices. It is also instructionally effective for a core task, as those who trained on MIST-VR were shown to be better at a laparoscopic cutting task than those who had not.²²

Our study intends to develop basic metrics, to evaluate the validity of the metrics, and to apply the metrics to a curriculum that is then instantiated into the endoscopic sinus surgery simulator (ES3). Here, we present our work in progress and discuss the preliminary results.

APPROACH

The first step in our investigation of improving patient safety with surgical simulation, and specifically for endoscopic sinus surgery, was to build an interdisciplinary and multi-institutional team with expertise in surgery, technology development, and training and evaluation methods. This team created a comprehensive curriculum related to endoscopic sinus surgery and simulation. An information data repository was created, allowing simulator performance data from various geographic sites to be uploaded and analysed via the internet. This database permits a wide range of analyses and standardised comparison of results from various subjects and control groups.

Endoscopic sinus surgery simulator

In response to the need for a stimulator to help train the novice sinus surgeon, defence contractor Lockheed Martin (spearheaded by senior otolaryngologist, Major Charles Edmond at Madigan Army Medical Center) developed an ESS training simulator, with funding from the US Army and the Telemedicine and Advanced Technology Research Center (TATRC).^{23 24} Using both visual and haptic (force) feedback to create a virtual reality environment, the ES3 was developed to teach core ESS procedures to otolaryngology residents.^{23–25}

Key box 1

- Simulation will potentially be integrated into surgical training.
- Our study intends to develop basic metrics, to evaluate the validity of the metrics, and to apply the metrics to a curriculum that is then instantiated in the endoscopic sinus surgery simulator (ES3).

The ES3 is composed of four principal components:

- a Silicon Graphics Incorporated computer which serves as the simulation host platform;
- a haptic system controller PC which performs the requisite high rate control of a physical instrument handle associated with a set of virtual surgical instruments;
- a virtual voice recognition instructor PC which responds to spoken commands controlling the simulator;
- an electro-mechanical platform which houses a physical replica of an endoscope, a mechanically linked surgical hand tool handle, and a mannequin of the external head anatomy.

The simulator allows the user to carry out endoscopic sinus surgery on a virtual patient using a wide range of simulated surgical instruments. The student surgeon explores the virtual anatomy by manipulating the simulated endoscope in the nose of a mannequin as they would in a real patient. With the help of the virtual instructor PC, the student is able to ask for various instruments audibly and have them appear on the monitor upon demand. The surgical field is lifelike and all instruments except the endoscope provide force (haptic) feedback through the universal instrument handle, allowing the student to feel tissue resistance appropriate to the tool (for example, resistance for scissors, vibration for a debrider).²⁶

As the students advance in their training on the simulator, they are given more complex tasks to perform, ranging from simple endoscopic navigation to an ethmoidectomy or sphenoidectomy in the presence of bleeding. The student starts training in an abstract novice environment. Here, the trainee performs representative abstract tasks (such as moving the endoscope through a series of hoops) and becomes familiar with the array of 23 surgical tools. This novice level training improves hand–eye coordination through an immersive experience that replicates the typical surgical interface. Upon achieving an appropriate score at this level, the student moves to the intermediate training level where abstract training aids (such as navigation hoops and anatomical labels) appear in the context of sinus anatomy, replicating the true surgical environment. The system prompts the student to carry out the appropriate tasks while learning the surgical procedure within this annotated anatomy. In the advanced level, the student performs full procedures without the benefit of the abstract training aids.

Throughout the simulation the virtual instructor points out mistakes, errors, and misses. The simulator records both overall and task specific scores for each student's performance. The overall score is a function of time to complete the task and task accuracy relative to optimal performance. Score penalties are also given for excessive time and for hazards hit (areas of the anatomy the surgeon should avoid, such as the optic nerve and lamina papyracea).

Metrics development

A conference, Metrics for Technical Skills, was convened to identify a set of objective measurements that could be assigned to the individual components of an endoscopic sinus operation. Before the conference, there had been no publication of a classification of errors in endoscopic sinus surgery, or on methodology for identifying how an error occurs, how to measure each error (metrics), or what outcomes should be reported. Experts from around the world were assembled, representing otolaryngologists, residents, anaesthesiologists, and the simulation and training field.

An initial workshop enumerated the metrics currently in the ES3 and suggested other possible metrics to be added. During this closed session, experts used a modified Delphi method of forecasting (Rand Corporation, 1969) to identify,

define, classify, and assign quantifiable measurements to generalisable errors that the ES3 should measure in order to provide a uniform framework for sharing, comparing, and analysing data with other investigators. At the same time, this workshop provided iterative feedback of the selected metrics based upon preliminary results using the ES3.²⁷

Validation of the metrics has been time consuming and very detailed. Careful study design and coordination across residency programmes is needed to obtain the statistical power to determine the face, construct, content, and concurrent and predictive validity of each metric. Correlations with previously validated tests of fundamental skills and psychomotor abilities on the MIST-VR, visuo-spatial, and pictorial surface orientation (PicSOr) were also assessed to strengthen our conclusions. The main hypotheses for metrics development are given below in the section on testing and validation studies.

Curriculum development

Keerl²⁸ showed that complications were reduced when surgeons underwent a multimedia learning programme before carrying out sinus surgery, showing that they had fewer dural and orbital complications. In order to standardise training on the ES3 (of surgeons, residents, and medical students alike) a curriculum was developed. By organising information to assist in training the user, regardless of education training level, JIU assisted by DR provided a refined state of the art ESS curriculum. The curriculum has been incorporated onto a multimedia compact disc and includes audiovisual information demonstrating both successful and erroneous operations at each training level, and providing a written background, anatomical information, and other helpful tools.

To maximise skill acquisition and gauge trainees' progress and patient outcomes, we further developed an objective training environment using two of the most sophisticated simulators developed to date. One of these is the ES3. The other is the MIST-VR (Mentice Corporation, Gothenburg, Sweden), which provides an abstract environment (using spheres, cubes, and virtual surgical instruments) that requires hand-eye coordination analogous to basic surgical manoeuvres. These two simulation devices have complementary objectives with minimum redundancy. The emphasis of the MIST-VR is on psychomotor skills and fundamental manual abilities, such as spatial accuracy, hand-eye coordination, ambidexterity, navigation, and three dimensional perception. The ES3 is a procedure simulator that trains and assesses task performance within the context of a surgical procedure (such as injection, which requires navigation, ambidexterity, and accuracy).

To gain a better understanding of the present state of resident training, we also conducted a survey of the major centres participating in this research project, by means of a questionnaire submitted to the otolaryngology programme chairs in the participant institutions. They were asked to describe the residents' preparation for ESS and the methods used to evaluate their readiness to undertake the procedure.

Testing and validation studies

To achieve the full potential of this technology, many subjects need to be tested to evaluate its efficacy and thereby iteratively modify the ES3 and other devices, as experience is obtained. This has required testing at various sites in distant geographical locations, as the number of otolaryngology residents in each programme is quite limited (as compared, for example, with general surgical residencies).

Discriminant and concurrent validity study

We first sought to demonstrate discriminant validity (that is, the lack of a relation among measures which theoretically

Key box 2

- The ES3 is one of the critical devices used in this research.
- Comprehensive metrics for an entire surgical procedure (endoscopic sinus surgery) have been developed.
- A standardised curriculum for the ES3 has been developed.

should not be related) and concurrent validity of the metrics on the ES3 (the system's ability to distinguish among groups that it should theoretically be able to distinguish), and to show that the simulator can accurately and reliably assess performance. Expert endoscopic sinus surgeons, otolaryngology residents, and novices (medical students) were tested individually on 10 separate occasions on each training level to establish performance parameters. Performance criterion levels from senior endoscopic sinus surgeons from across the country were obtained and the variability of these scores determined. This research is unique in that no other benchmark criteria for sinus surgery have ever been accumulated. Differences among these groups may not simply be in average performance score but also in performance variability and in opportunities for safety improvement.

All subjects were given a detailed demonstration of the ES3, a description of the different instruments and their function, as well as the multimedia CD described above. They were then surveyed using a five point Likert scale that elicited their perceptions of the quality and appropriateness of various aspects of the simulation and curriculum. Negative feedback from this group (such as on the various weightings in the scoring algorithm) was acted upon where appropriate and attainable. This provided a broad based assessment of the validity of the ES3 content.

We have used our web based database archive to compile subject data from Montefiore Medical Center, Madigan Army Medical Center, and the US Naval Medical Center in San Diego. Incorporation of new attending, resident, and student data is ongoing.

Construct validity study

We hypothesised that student training on the ES3 would also demonstrate construct validity. This means that the simulator reliably measures psychomotor, visuo-spatial, and perceptual abilities, and correlates positively with objective tests of such fundamental abilities that have already been shown to predict surgical performance. Gallagher's group demonstrated the functional involvement of psychomotor ability in the adaptation, consolidation, and development of skills in endoscopic surgery.²⁹ To date no general task measuring an individual's pure psychomotor ability has been assessed as a possible prospective metric for endoscopic surgery.

Thirty four medical students and four otolaryngology residents from the Albert Einstein College of Medicine were assigned to perform one trial on the ES3 at the novice level, six tasks on the MIST-VR at the intermediate level, three visuo-spatial tasks, and one set of 35 trials on PicSOr figures. The procedure, the use of these devices, and their features are as follows.

Measurement of psychomotor ability

This was assessed through the MIST-VR system, which has been independently validated as a measure of endoscopic psychomotor performance. The subject executes specific tasks

that are functionally related to tasks carried out in laparoscopic surgery and then receives feedback about performance.^{30, 31} The subject simulates grasping tissue, transferring it from one gripper to the other, running the bowel by using hand over hand transfer, removing a tool from the operating field and reinserting it accurately, cauterising three sub-targets, and maintaining objects within the target box while cauterising three consecutive subtargets.³²

Visuo-spatial ability

This was assessed using the card rotation, cube comparison, and map planning tests from the kit of factor referenced cognitive tests generated by the Educational Testing Service.³³ These tests assess the subject's appreciation of the spatial representation of objects that are arranged in various ways.

Perceptual assessment

This was measured through a test called pictorial surface orientation (PicSOR) based on the techniques described by Cowie.³⁴ Each item is a picture on a computer monitor, showing a spinning arrowhead with its point touching the surface of a cube or a sphere. The subject manoeuvres the arrowhead (using cursor keys) until its shaft is perpendicular to the object's surface at the point where they touch. This is a relatively pure test of a subject's ability to recover the pictorial cues that specify orientation of structures in (virtual) pictorial space, and to compare the implied orientations. The most important measure of performance is the correlation between theoretically correct arrowhead orientation and the setting chosen by the subject, and the slope of the fitted regression line.

Predictive validity study

Does ES3 training generalise from VR to the operating room? Establishment of the reliability and validity of the simulation system is of crucial scientific importance. However, for the endoscopy surgeons, residents, and patients, the most important test is whether training in a virtual environment will transfer or generalise to the real world of the operating room. Experience from the aviation training field clearly suggests that it will. However, merely assuming that this will happen is insufficient. The purpose of this study is to demonstrate transfer of training objectively, addressing directly the question of whether training on the ES3 will improve the surgical performance of otolaryngology residents in actual surgery and reduce surgical errors. Statistical power for this study requires the collection of appropriate resident data from multiple institutions.

Otolaryngology residents in their junior years are included both in the experimental group (those who receive conventional sinus surgery training as well as the ES3 training curriculum) and in a control group (those who receive only conventional sinus surgery training). Following training, one of their first five sinus surgery procedures is videotaped for all individuals within each group, and this anonymous taped procedure is then rated by a select group of trained raters. Residents from Albert Einstein College of Medicine, New York University, New York Eye and Ear Infirmary, Madigan Army Medical Center, and the US Naval Medical Center in San Diego are participating. Control group residents are from the otolaryngology programmes at Yale University, Mount Sinai Hospital (New York), and Columbia-Presbyterian Medical Center.

Web database development

The database is the fundamental unit that integrates the project. The metrics component identified quantifiable measures which then become the fields for the database. The ES3 acquires measurements during training and can submit data in an automated and standardised format to the

Key box 3

- Concurrent and discriminant validity of the ES3 has been established.
- To date no general task measuring an individual's pure psychomotor ability has been assessed as a possible prospective metric for endoscopic surgery.
- ES3 has demonstrated construct validity.
- The purpose of this study is to demonstrate objectively a reduction of errors and transfer of training from the ES3 curriculum to surgical performance of otolaryngology residents in actual surgery.

central database for storage, analysis, and information exchange among the participating institutions. Additionally, the database includes previous University of Washington human interface technology (HIT) laboratory validation studies on the ES3.¹⁸ The centralised web based database also provides an analysis software tool set and generates reports that are used to improve the project iteratively, as well as reports for outcomes analysis. Thus it permits the standardisation of the effort on a national scale and provides a single resource responsible for the security and integrity of the data.

Both simulator performance and comparison studies of actual surgery are being incorporated and will be used to guide further development of the ES3, coupling the metrics of surgical simulation with curriculum development and training, objective analysis of surgical performance, and surgical science. This database will be compatible with regional and national surgical simulation training data repositories, which are currently being considered by professional and national credentialing organisations.³⁵

Authorised researchers can select and analyse datasets and generate objective scores for surgical skill and safe procedure execution. The database allows the research team to carry out the following:

- Quantitatively configure and validate the ES3 through comparison with measurements of the same operation on inanimate models, and eventually animals and humans.
- Acquire data from real experiments, surgical simulators, and eventually surgical telerobots in a uniform format. These data can be evaluated for internal validity and consistency, and linked with other, standardised measures of cognitive and psychomotor skill for individual student surgeons.
- Quantitatively define types of error based on the metrics specified during the metrics workshops and during the development of the simulator, to include:
 - incorrect manoeuvres, with violation of tissue or instrument tolerances;
 - correctly performed instrument manoeuvres that are out of sequence or inappropriate for that part of an operation;
 - inefficient force patterns or application, and inefficient manoeuvres or sequences of manoeuvres;
 - inappropriate variability in technical performance;
 - inappropriate "dwell time" or "lack of progress," indicating indecision or confusion.
- Provide real time, contextually accurate analysis and feedback to the student surgeon, for error recognition and correction, in addition to objective comparison

(statistical similarity) with other subjects already represented in the database.

- Provide outcomes that represent an overall assessment of technical skill for an individual surgeon. These statistics, when assessed together with other archived measures of cognitive and interpersonal skill, should also provide a first order metric for the global assessment of competency.
- Provide data in a near real time, recursive, and iterative feedback cycle to support curriculum development.
- Acquire datasets for new operations and create models based on them.
- Support analysis—based on demographics, training, and performance—across many simulated procedures or groups of surgeons to define parameters of competency, skills, and training for credentialing, regulatory, and policy purposes to appropriate surgical boards and societies, and to state and federal agencies.

FINDINGS

Metrics development

Errors in ESS were identified, and each error was classified into taxonomy according to type (technical—associated with manual skills; cognitive—resulting from mistakes in the conscious decision making process, the remembering of procedural details, or both; or combined).²⁷ The list was approved by consensus, and quantifiable measurements were then assigned to each error. Some errors, such as “wrong tool choice” or “contact with wall” were simple “yes or no” measures. Other errors, such as “injury to the lachrymal duct” or “too much scope rotation” have clear and measurable end points. Some errors, such as “past pointing” or “lack of perspective”, may need further quantification and will be evaluated when the video recordings of each procedure are reviewed by the rater panel (with a criterion of greater than 80% inter-rater reliability).

Lastly, outcome measures were then derived and agreed to, in the broad areas of technical, cognitive, and combined skills, as well as measures for global assessment. The report of the Consensus Workshop on Errors has been vetted by the American Academy of Otolaryngology–Head and Neck Surgery and by the American Rhinologic Society and is considered as a standard. This was the first time that metrics have been assigned to an entire surgical procedure of any kind.²⁷ It forms the basis of our scoring procedure for surgery

undertaken on patients during the predictive validation study.

Curriculum development

This effort has resulted in a refined, state of the art curriculum equipped with both audio and video interactive information. Preliminary results show that this CD curriculum greatly facilitates the trainees’ compliance with the study protocols and allows us to help eliminate certain levels of bias caused by the different educational starting points of our subject participants. We view this teaching manual and disc as an essential part of the ES3 educational experience and have further incorporated the cognitive enhancement components of our research.

The residency programme survey yielded the following results: first, residents often do not begin doing endoscopic sinus surgery until their last year of otolaryngologic training; second, only a little more than half the institutions that responded gave formal lectures before residents were allowed to perform ESS, and the number of hours of lectures also varied by institution; third, only a little more than half the programmes required completion of a reading list before residents were allowed to perform in the operating room; and fourth, most residency training programmes afforded their residents an opportunity to practice their ESS techniques on cadaver heads before their operating room experience.

While our sample of ESS training programmes is small, it represents a broad range (that is, private and public universities, civilian and military medical centres, distributed geographically across the USA). More extensive sampling is currently under way to verify these trends and to identify any consistent differences across these classes of programme.

Testing and validation studies

Discriminant and concurrent validity study

Assessment of the data clearly shows that there are significant group differences in the performance of medical students, residents, and attendings on the ES3 (fig 1). As one would expect from a simulator that truly reflects sinus surgery, attendings perform the best, with residents close behind. Although medical students improve their performance over the course of the study, they perform substantially below par with the residents, with par being the performance level of the attendings. Given the number of participants that we have enrolled from all the study sites, these results bear substantial statistical power.

An important observation from our study of medical students’ performance is that, after a critical period of training has been accomplished, the performance of the trainee does not diminish substantially despite a period of absence from simulator training as great as one month. These training gaps ranged from 11 to 90 days (average 34.8 days).³⁶

Construct validity study

Results from this study have shown conclusively that there are several correlations among these many validated tests and the ES3. For example, we have found through a simple regression model that scores from the picture surface orientation test (a validated instrument that was developed to identify aptitude for the recovery of three dimensional structures from two dimensional images) correlate highly with the hazard score on the ES3 (the score that represents the number and severity of errors that are committed during a simulated surgical procedure).

One of the strongest correlations observed was between the total MIST-VR score for all six tasks and the overall trial score on the ES3. When the PicSO perceptual task, the visuo-spatial cube comparison task, and the MIST-VR psychomotor task total scores were included in a multiple regression

Key box 4

- The web based intelligent information interface and database is the fundamental unit that integrates the project.
- Storage, analysis, and information exchange through the internet among the participating institutions is essential.
- This database will be compatible with regional and national surgical simulation training data repositories.
- The database will provide real time, contextually accurate analysis and feedback to the student surgeon, for error recognition and correction, in addition to objective comparison (statistical similarity) to other subjects already represented in the database.
- Parameters of competency, skills, and training for credentialing, regulatory, and policy purposes are being defined.

Figure 1 Endoscopic sinus surgery simulator (ES3) performance curves over trials for attendings, residents, and medical students.

model, they were found to be strong and statistically significant predictors of ES3 performance.

Predictive validity study

Earlier pilot data from the original Madigan validation studies suggest that some aspects of surgical skill will show definitive transfer of training, while others may not. Because of the complexity of the current study and the need to enlist residents in a particular year of their residency, the vast majority of the data collection for this study is still ongoing, and preliminary results were not available at the time of writing.

Web based database

The four major components in our database architecture are, first, the ES3 simulator, currently running on an SGI computer which will remain isolated from direct internet access; second, the proctor's software, running on a PC that communicates directly with the SGI through a local network and with the central database server via the internet; third, the central SQL server, housed at the University of Washington, which manages database input and user access functions; and fourth, the database access software for the general public and project investigators, using a standard web browser interface over the internet.

We modelled the look and feel of our web site after the existing Agency for Healthcare Research and Quality (AHRQ) M&M site (<webmm.ahrq.gov>). All users enter the site through the public face page and then log on to gain access to additional functions. Three methods of accessing the simulator trial database are supported:

- immediate access to the trainee's score and learning curve, relative to appropriate comparison groups (for example, senior residents), is available through the proctor's interface during training trial debriefing;
- simple descriptive statistics and data graphics for user selected subsets of the database are available to project investigators directly from the web site;
- user selected data subsets can be downloaded by investigators for further statistical analysis on their local machine (using standard data analysis software, such as Excel or SPSS).

In addition to the simulator trial data, the server archives and provides access to various project documents including the ES3 user manual, various instructional video clips, protocol descriptions, graphics, and drafts of working papers.

The public portion of the site provides access to study descriptions, published reports, a glossary of terms related to surgical training and simulation, and pointers to related simulation and patient safety websites. The central database accepts data from each simulator's local proctor machine over the Internet through a SOAP based interface. The public portion of the database website can be explored at <<http://utopia.hitl.washington.edu/ahrq/>>. Access to the simulator trial data and the web based analysis functions is currently restricted to project personnel.

DISCUSSION

Curriculum

As a result of the rigorous process for deriving objective measurements for endoscopic sinus surgery, this research introduces an emphasis on error reduction as a formal component of surgical training. Students were taught precisely what the errors are and how to avoid them—in a uniform and comprehensive manner. An interesting observation is that this approach has instilled a unique “culture of safety” into the participating surgical training programmes.

Training to a proficiency level (that is, until competent)—whereby no student operates upon a patient until they have been objectively “proven” to be safe—may fundamentally disrupt and reshape surgical training. Today surgeons train for a fixed period of time; future surgeons may have a variable (shorter or longer) residency programme, depending upon how quickly they attain competence. What this guarantees is that, no matter how long the individual resident's training programme lasts, any patient operated upon by that surgeon will have a safe surgical procedure. One additional benefit of this changing perspective may be to provide a method of accommodating to the new limitation on resident work hours. It is currently uncertain how it will be possible to provide enough training and to verify the quality of the training. Focused learner specific simulation training may provide a more efficient route to surgical proficiency and better ensure patient safety throughout the surgical residency.

Based upon our survey, we conclude that there is clearly a need to develop a curriculum content that is systematic, uniform, comprehensive, and well structured to include cognitive content. In the particular case of ESS, a validated curriculum to increase the effectiveness and safety of motor skills is also absolutely essential. We continue to hypothesise that such a structured curriculum will improve effectiveness, reduce errors, and improve safety in participants compared with control subjects who learn by traditional methods.

Discriminant and concurrent validity

Because of the substantial number of medical students who have been trained, our numbers hold statistical significance and power. We feel that this firmly establishes the discriminant and concurrent validity of the ES3 and provides benchmark criteria for performance on the ES3 to which other trainees can be compared. The conclusions based upon these data have been presented at the American Academy of Otolaryngology–Head and Neck Surgery Annual Meeting in San Diego, California (September, 2002) and at the Medicine Meets Virtual Reality conference in Newport Beach, California (January, 2003). We also conclude that the ES3 affords training in a complicated skill such that it produces long lasting and memorable performance pathways.³⁶

Construct validity

Scores on the ES3 correlate positively with scores on other previously standardised measures of psychomotor, perceptual, and visuo-spatial skill. Task scores from the MIST-VR trainer correlate strongly with dissection scores, navigation

scores, and injection scores on the ES3, all of which have strong dependence on hand–eye coordination. The ES3 accurately provides an assessment of the technical skills that it proposes to measure, thus providing evidence for construct validity of this instrument. This study suggests that individual students who have difficulty extracting three dimensional information from two dimensional images appear to have equal difficulty manipulating a three dimensional simulated patient using a two dimensional monitor. Performance in all of these tests accurately predicts performance in the ES3.

Web database

The creation of the interactive web based database is a core component of our research programme. Without it, the acquisition and integration of data from remote ES3 sites would be tremendously difficult. Remote ES3s located on various campuses are necessary for accumulating adequate numbers of subjects in a field such as otolaryngology, given that residency programmes are so small. In addition, the ES3 database will serve as a prototype for the creation of a national database whereby surgical skills can be compared among subjects to an established criterion. No other repository for surgical skills exists to date. Our database design is not tied specifically to the ES3, but is readily generalisable to other surgery (and non-surgical) simulators; we envision an AHRQ sponsored resource that will be useful across medical disciplines where simulation is used in resident training.

Consortium

We have been able to establish a working team of experts in the field of surgical simulation, with consultants from other critical fields already immersed in a similar training approach (such as the flight industry). The primary simulator used in this study (ES3) will be further enhanced and made more robust and sophisticated through the collaboration of industry leaders. The use of surgical simulation will undoubtedly be one of the most critical subjects in surgical training, performance evaluation, certification, and ultimately patient safety.

Limitations

The endoscopic sinus surgery curriculum would benefit from being based on-line, and we are trying to establish this. The participation of subjects (and controls) is limited by the small number of residents within any otolaryngology training programme. We are therefore trying to recruit as many programmes as possible. The ES3 itself was developed seven years ago and we await the next version. We anticipate that this will occur with the projected commercialisation of this device.

Conclusion

Our research involves a revolutionary form of training in endoscopic sinus surgery. We are establishing a training methodology that can reduce *all* surgical errors by incorporating simulation into a rigorous curriculum that is extensible, scalable, and generalisable. We plan to demonstrate that the approach is *extensible* by adding other endoscopic procedures that can be carried out on the ES3 platform. We will demonstrate that it is *scalable* by increasing the number of centres and students that can be archived and analysed through the web based database. Finally, we will show *generalisability* by incorporating the methodology and core curriculum into non-otolaryngology training programmes.

This work is supported by grant No 1 R18 HS11866–01 of the Agency for Healthcare Research and Quality.

Authors' affiliations

M P Fried, Department of Otolaryngology, Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, New York, USA
R Satava, Department of Surgery University of Washington, Seattle, Washington, USA
S Weghorst, Human Interface Technology Laboratory, University of Washington
A G Gallagher, Emory Endosurgery Unit, Emory University School of Medicine, Atlanta, Georgia, USA
C Sasaki, Department of Surgery Division of Otolaryngology, Yale University School of Medicine, New Haven, Connecticut, USA
D Ross, Department of Surgery, Division of Otolaryngology, Yale University School of Medicine
M Sinanan, Department of Surgery, University of Washington
J I Uribe, Otolaryngology Surgery Simulator Center, Montefiore Medical Center, Bronx, New York
M Zeltan, Otolaryngology Surgery Simulator Center, Montefiore Medical Center
H Arora, Albert Einstein College of Medicine, Bronx, New York
H Cuellar, Otolaryngology Surgery Simulator Center, Montefiore Medical Center

REFERENCES

- Hudgins PA. Complications of endoscopic sinus surgery. *Radiol Clin North Am* 1987;**31**:21–2.
- Stankiewicz JA. Complications of endoscopic intranasal ethmoidectomy. *Laryngoscope* 1987;**97**:1270–73.
- Tendick F, Jennings RW, Tharp G, et al. Sensing and manipulation problems in endoscopic surgery: experiment, analysis and observation. *Presence* 1993;**2**:66–81.
- Pelausa EO, Smith K, Dempsey I. Orbital complications of functional endoscopic sinus surgery. *J Otolaryngol* 1995;**24**:154–9.
- Rosser JC, Herman B, Risucci DA, et al. Effectiveness of a CD-ROM multimedia tutorial in transferring cognitive knowledge essential for laparoscopic skill training. *Am J Surg* 2000;**179**:320–4.
- Reznick RK. Teaching and testing technical skills. *Am J Surg* 1993;**165**:358–62.
- Gorman PJ, Meier AH, Krummel TM. Simulation and virtual reality in surgical education: real or unreal? *Arch Surg* 1999;**134**:1203–8.
- O'Toole RV, Plyter RR, Krummel TM, et al. Measuring and developing suturing technique with a virtual reality simulator. *J Am Coll Surg* 1999;**189**:114–27.
- Satava RM. Virtual reality surgical simulator. The first steps. *Surg Endosc* 1993;**7**:203–5.
- Gorman PJ, Meier AH, Krummel TM. Computer-assisted training and learning in surgery. *Comput Aided Surg* 2000;**5**:120–30.
- Peugnet F, Dubois P, Rouland JF. Virtual reality versus conventional training in retinal photocoagulation: a first clinical assessment. *Comput Aided Surg* 1998;**3**:20–6.
- Gladstone HB, Raugi GJ, Berg D, et al. Virtual reality for dermatologic surgery: virtually a reality in the 21st century. *J Am Acad Dermatol* 2000;**42**:106–12.
- McGovern KT. Applications of virtual reality to surgery. *BMJ* 1994;**308**:1054–5.
- Satava RM. Virtual endoscopy: diagnosis using 3-D visualization and virtual representation. *Surg Endosc* 1996;**10**:173–4.
- Baillie J, Evangelou H, Jowell P, et al. The future of endoscopy simulation: a Duke perspective. *Endoscopy* 1992;**24**:542–3.
- Fried MP, Moharir VM, Shinmoto H, et al. Virtual laryngoscopy. *Ann Otol Rhinol Laryngol* 1999;**108**:221–6.
- McCarthy A, Harley P, Smallwood R. Virtual arthroscopy training: do the virtual skills developed match the real skills required? *Stud Health Technol Inform* 1999;**62**:221–7.
- Weghorst S, Airola C, Oppenheimer P. Validation of the Madigan ESS simulator. *Stud Health Technol Inform* 1998;**50**:399–405.
- Chaudry A, Sutton C, Wood J, et al. Learning rate for laparoscopic surgical skills on MIST-VR, a virtual reality simulator: quality of human-computer interface. *Ann R Coll Surg Engl* 1999;**81**:281–6.
- Smith S, Wan A, Taffinder N, et al. Early experience and validation work with procedicus VA – the prosovia virtual reality shoulder arthroscopy trainer. *Stud Health Technol Inform* 1999;**62**:337–43.
- Prystowsky JB, Regehr G, Rogers DA, et al. A virtual reality module for intravenous catheter placement. *Am J Surg* 1999;**177**:171–5.
- Gallagher AG, McClure N, McGuigan J, et al. Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer virtual reality (MIST-VR). *Endoscopy* 1999;**31**:310–13.
- Rudman DT, Stredney D, Sessanna D, et al. Functional endoscopic sinus surgery training simulator. *Laryngoscope* 1998;**108**:1643–7.
- Edmond CV, Heskamp D, Sluis D, et al. ENT endoscopic surgical training simulator. In: Morgan KS, eds. *Medicine meets virtual reality*. Amsterdam: IOS Press, 1997:518–28.
- Wiet GJ, Yagel R, Stredney D, et al. A volumetric approach to virtual simulation of functional endoscopic sinus surgery. *Stud Health Technol Inform* 1997;**39**:167–79.
- Edmond CV, Heskamp D, Sluis D, et al. ENT endoscopic surgical training simulator. *Stud Health Technol Inform* 1997;**39**:518–28.

- 27 **Satava RM**, Fried MP. A methodology for objective assessment of errors: an example using an endoscopic sinus surgery simulator. *Otolaryngol Clin North Am* 2002;**35**:1289–301.
- 28 **Keerl R**. Value of multi-media educational software in training of the paranasal sinus surgeon. *Laryngorhinootologie* 2000;**79**:34–8.
- 29 **Gallagher AG**, McClure N, McGuigan J, et al. An ergonomic analysis of the fulcrum effect in acquisition of endoscopic skills. *Endoscopy* 1998;**30**:617–20.
- 30 **Gallagher AG**, McGuigan J, Ritchie K, et al. Objective psychomotor assessment of senior, junior and novice laparoscopists with virtual reality. *World J Surg* 2001;**25**:1478–83.
- 31 **Gallagher AG**, Satava RM. Objective assessment of experience, junior and novice laparoscopic performance with virtual reality: learning curves and reliability measures. *Surg Endosc* 2002;**16**:1746–52.
- 32 **Hamilton EC**, Scott DJ, Fleming JB, et al. Comparison of video trainer and virtual reality training systems on acquisition of laparoscopic skills. *Surg Endosc* 2002;**16**:406–11.
- 33 **Ekstrom RB**, French JW, Harman HH, et al. Manual for kit of factor-referenced cognitive tests. Princeton, New Jersey: Educational Testing Service, 1976.
- 34 **Cowie RID**. Measuring and modeling of perceived slant in surfaces represented by freely viewed line drawings. *Perception* 1998;**27**:505–40.
- 35 **Nahrwold D**. The competence movement: a report on the activities of the American Board of Medical Specialties. *Bull Am Coll Surg* 2000;**85**:14–18.
- 36 **Uribe JL**, Ralph WM, Glaser AY, et al. Learning curves, acquisition and retention of skills trained with the endoscopic sinus surgery simulator. *Am J Rhinol* 2004;**18**:87–92.