VIRTUAL REALITY SIMULATIONS AND INTERVENTIONAL RADIOLOGY

MAX BERRY

UNIVERSITAS REGIA GOTHOBURGENSIS

GÖTEBORG

2007

Printed at Vasastadens Bokbinderi AB, Göteborg, Sweden

ISBN: 978-91-628-7097-3

VIRTUAL REALITY SIMULATIONS AND INTERVENTIONAL RADIOLOGY

MAX BERRY

DEPARTMENT OF RADIOLOGY

INSTITUTE OF CLINICAL SCIENCES

THE SAHLGRENSKA ACADEMY AT GÖTEBORG UNIVERSITY

GÖTEBORG, SWEDEN

2007

To:

SIDSEL

ABSTRACT

INTRODUCTION: Use of virtual reality (VR) simulators in endovascular interventional education has become increasingly popular yet many questions surrounding this nascent technology remain unanswered. While progress has been made in other disciplines such as endoscopy and minimally invasive surgery, scientific evidence investigating endovascular simulations remains limited. The general aim of this dissertation was to conduct validation studies to elucidate the potential for skills acquisition and assessment outside of the catheterization laboratory using VR simulation. Endovascular skills transfer from VR-Lab to the porcine laboratory (P-Lab) was also investigated. An economic analysis was performed to assist in the establishment of a realistic VR implementation strategy. **MATERIALS AND METHODS**: Simulator validations were conducted by comparing performance metrics collected from novices and experienced physicians using Student's t-test. Performance metrics were recorded by the simulator while participants treated simulated patients suffering from renal artery stenosis (RAS) and carotid artery stenosis (CAS). Endovascular skills transfer was tested using the P-Lab as an approximation of the human catheterization laboratory. A group of endovascular novices were evaluated in the P-Lab and the VR-Lab using an objective skills assessment of technical skills (OSATS), yielding a Total Score. Participants were then randomized into different training groups, put through their assigned training schema and subsequently re-evaluated in both laboratories. ANCOVA analysis was conducted to compare the cumulative effect each type of training had on Total Score. Consumable and rental fees from the skills transfer study were used to calculate the comparison data for the economical analysis. **RESULTS**: Face validity was demonstrated for both the renal and carotid artery stenosis modules. Neither construct validity study produced results which differentiated between the expert and novice performance metrics except for fluoroscopic and procedural times. VR-Lab training sessions generated skills which improved P-Lab performances. VR-Lab training cost less than the P-Lab using our economical analysis. **CONCLUSIONS**: Despite demonstrating face validity, VR-Lab simulations should not be used alone for skills assessment outside of the catheterization laboratory in its present form. Skills learned in virtual reality transfer favorably to the P-Lab and simulation training seems to offer a viable alternative of non-clinical training. The VR-Lab affords a more economical method to teach and practice endovascular skills compared to the P-lab. Further research is needed to elucidate the relative efficacies of both training methods.

ORIGINAL PAPERS

This dissertation is based upon the following articles:

I. Berry M, Lystig T, Reznick RK and Lönn L. **Assessment of a Virtual Interventional Simulation Trainer**. Journal of Endovascular Therapy, Apr 2006; 13(2), 237-43.[*]

II. Berry M, Lystig T, Beard J, Klingenstierna H, Reznick RK and Lönn L. **Porcine Transfer Study: Virtual reality simulator training compared to porcine training in endovascular novices**. Cardiovascular and Interventional Radiology, May-June 2007; 30(3), In press.[**]

III. Berry M, Hellström M, Göthlin J, Reznick RK and Lönn L. **Endovascular Training using Animals or Virtual Reality Systems: An Economic Analysis**. Submitted.

IV. Berry M, Lystig T, Reznick RK and Lönn L. **The Use of Virtual Reality for Training Carotid Artery Stenting: A Construct Validation Study**. Submitted.

[*] Reprinted with kind permission from the Journal of Endovascular Therapy ©International Society of Endovascular Specialists

[**] Reprinted with kind permission from Cardiovascular and Interventional Radiology published by Springer Science and Business Media.

CONTENTS

BACKGROUND

ENDOVASCULAR INTERVENTION

Endovascular intervention is a highly specialized and potentially dangerous procedure performed by interventional radiologists and other specialists. Currently available treatment options include percutaneous transluminal angioplasty, stenting, thrombolysis, embolization, intravascular filters, plaque excision and foreign body removal. Traditionally, basic catheterization skills were acquired by radiology residents—as well as cardiology and vascular surgical fellows—while performing routine angiograms in the catheterization laboratory (Cath-Lab) under the close supervision of an experienced mentor. The practical skills gained during these diagnostic procedures were directly transferable to the execution of endovascular interventions. The traditional minimal number of digital subtraction angiographies (DSA) needed prior to advancing on to endovascular interventions ranged from 50-100 although the actual number varied by location, sub-specialty and procedure.[1-5] However, improved diagnostic imaging modalities such as computed tomography- (CTA) or magnetic resonance- (MRA) angiography combined with a more critical view regarding the acceptability of training on patients has led to a decrease in the amount of angiograms, constituting an ethical dilemma.[6-9] Both factors have jeopardized the normal training arena for interventional fellows.

In response, many institutions have turned to non-clinical training methods such as animal models, typically adult swine in a pig laboratory (P-Lab), to assist endovascular fellows in making the jump from routine angiography to interventional techniques.[10-14] Such a strategy acknowledges the technical difficulties and the early learning curves experienced during transitioning.[2,3,15,16] The goal is to provide the trainee a place to learn and practice which simulates the actual catheterization laboratory while removing unnecessary patient risk.[17,18] Excluding ethics and availability of animals for now, one can confidently state that healthy pigs are not an accurate simulation of pathologic human anatomy. Nonetheless, it is *in vivo* and allows the use of real instruments to practice for safe interventions.

Another solution to these conflicting demands is the *ex vivo* use of the virtual reality laboratory (VR-Lab). The rapid advances in computer technology over the last decade has heralded phenomenal technology such as full procedure virtual reality simulators capable of reproducing operative experiences outside of the catheterization laboratory.[19] VR simulators can systematically expose the trainee to a broad variety of pathologies in a safe, reproducible environment thereby eliminating the usual randomness of training.[20-22] Other unique advantages available with VR technology, such as objective feedback and the removal of the need for continual direct supervision have already been identified by other research.[20,23] However, in their current forms, endovascular simulators require introduction to the machine itself and close supervision during training to obtain the maximum benefit and prevent the development of dangerous habits. These systems allow endovascular fellows to develop basic psychomotor and procedural skills and fully acquaint themselves with the clinical environment prior to traditional patient based mentoring. Some proactive specialties have recommended evaluation, piloting and introduction of VR training into their specialty training to take advantage of VR's unique capabilities.[14,24,25] Indeed, the surgical fields of endoscopy and laparoscopy have seen an explosion of research examining the usefulness and validity of VR simulations which is too numerous to list.

LEARNING THEORY

Cognitive learning—defined as mental couplings between knowledge and physical skills—states that learning occurs when humans attain equilibrium between their reactions and their surroundings through assimilation and accommodation.[26] Assimilation, according to Piaget, is simply learning by addition. In other words, as one is exposed to new stimuli, a repertoire of experiences is added to the existing knowledge base. Accommodation occurs when new situations are encountered which require the use of pre-existing knowledge. This process, although not directly creating new knowledge, entails the deconstruction and rebuilding of existing knowledge into a form which is applicable to the alien situation. Accommodation presents more difficulty for the learner yet offers a deeper understanding. Nissen expanded this line of thinking with his notion of cumulative learning, which he defined as the use of all of ones cognitive schemas applied to a completely new learning environment—creating a new set of knowledge.[27]

Psychodynamic learning, as proposed by Vroom, tells us that all learning is dependent on feelings and motivations.[28] Thus the feelings, both positive and negative, we experience during a learning situation are an integral part of the process and affect the outcome. Motivation to learn falls within this category as well. A learner who believes that they can master a new topic which will prove useful in a fashion that improves their own situation stands a much higher chance of succeeding.[27]

Bruner's societal learning dimensions focus on the fact that everything a person does is influenced by, and in turn influences, a cultural and social context.[27] Societal learning is thus an interaction between people, language and objects and is dependent upon social and contextual influences. A more concrete example of this type of learning is the mentor-apprentice model.[29] The apprentice begins his studies as a passive observer who, over time and in various contextually specific situations, takes on increasing degrees of responsibility under the guidance of a master. This very model is the one which has dominated surgical and interventional radiology skills teaching and continues to be the predominant method to date.

MOTOR SKILL DEVELOPMENT

Behavioral psychologists divide motor skills learning into three distinct levels.[30] Psychomotor skills are those which, after numerous repetitions, may be partially or fully automated by the motor cortex to a level of unconsciousness.[31] A common example is the ability to ride a bicycle. Enormous amounts of energy must be used to coordinate muscle efforts to attain smooth, steady cycling at the beginning of the learning curve. Once these processes are internalized, one hardly needs to think about the actions. Instead, the actions are automated freeing larger amounts of working memory for other endeavors. Basic catheter manipulation skills in the endovascular suit fall into this skill category.

Procedural skills entail the learning of rules and/or steps. A practical example would be a cake recipe with instructions. One cannot simply add all ingredients into a large dish, throw it into the oven and expect a three-layered butter-cream masterpiece. Specific steps must be followed in a particular order to achieve the desired outcome. Similar logic applies in the human catheterization laboratory. One cannot hope to put a patient, the correct interventional tools and a novice into a room and expect satisfactory results. The procedure must proceed according to protocol if the patient is to benefit from treatment.

Lastly, cognitive skills encompass decision-making and feats of manual dexterity when faced with an unfamiliar or unexpected environment. In short, these are reactions to a new situation that are created and executed to attain a desired result based upon that particular individual's body of knowledge. This skill answers the question: Given an unknown, what would one do next? In surgical specialties, a novice's ability to correctly react to sudden, intraoperative complications is a defining step toward mastering their trade.

ENDOVASCULAR SKILLS

Although the exact definitions of what constitutes a skilled interventionalist remains unwritten, a behavioral psychologist approach seems most appropriate.[30,32] Psychomotor skill acquisition, sometimes termed generation, is the process of initial learning much like learning to ride a bicycle. During these fragile first steps, one needs to focus large amounts of concentration and expend tremendous effort to force the body to make the new motions in an appropriated fashion as instructed. Acquisition of the new psychomotor skill is a separate entity from practicing that skill until mastery.[19] The time from initial learning until mastery is usually referred to as the *learning curve* and, unfortunately, represents the period of time when most surgical errors are likely to happen.[2,33,34]

Traditional surgical skills allow the natural use of the senses to see, touch and smell as the maneuvers are being attempted. However, endovascular intervention robs the surgeon of the ability to see in three dimensions and to make use of the direct tactile feedback, known as haptics. Two dimensional fluoroscopy replaces the open wound and the interventional haptic feedback is miniscule by comparison. Thus, interventional psychomotor skills cannot be considered interchangeable with open surgical skills or laparoscopic ones.

Procedural skills represent the ability to follow a given procedures protocol closely. In essence, it is the ordered steps needed to perform a dance properly. Due to the nature of interventional techniques, e.g. singular femoral artery access, the methodical introduction of equipment through the arteriotomy according to protocol allows successful results. Failure to do so may render hours of preparatory work meaningless as the necessary instrument may not be able to be deployed properly causing lost Cath-Lab time, wasted equipment and extended risk exposure for the patient. Literally speaking, there is no room for error.

Lastly, cognitive behavior denotes how a person reacts, based on their inherent body of knowledge, when they meet with the unexpected. Simply stated, cognitive skills come into play when surgical difficulties or complications present themselves. For example, the textbook anatomy expected may actually be an unrecognizable congenital anomaly which forces the interventionalists to find another vascular route to the target site impromptu. Cognitive skills can only be fully developed in a real catheterization laboratory, human or otherwise, because, although it may one day be possible to incorporate such scenarios into the simulation milieu, the extent of unpredictability found in a real cath-lab includes too many variables to be included. Thus, while procedural complications, e.g. angioplastic balloon rupture, could be included into the simulator, human factors causing disturbances in work flow such as a non-cooperative patient who refuses to lay still during angiography or experiences nausea and vomiting during the most critical point in a procedure may not be as readily simulated.

Skills Assessment

Who, when and what should be the final judge of a fellow's readiness to enter the catheterization laboratory? Three paths are available. Traditionally, subjective approval from one's mentor marked the passage from apprenticeship to journeyman, i.e. independence in the Cath-Lab. While rooted in tradition, the major drawbacks of this method are its continued use of patients to train and dependence upon a random exposure to procedures. Such a practice is especially questionable when examined in the light of ethics-based medicine and the increasingly litigious atmosphere in which we work.

In contrast, validated metric based virtual reality assessments offer a completely objective, standardized model wherein fellows might attain metric benchmarks prior to independent catheterization laboratory work.[19,35,36] This seems promising, but places the grave decision-making responsibility on microprocessors incapable of human subtleties. Additionally, this requires reproducible metric validations for each parameter measured by the simulator, for each module and for every type of simulator.[37] Lastly, simulator metrics assesses virtual reality skills, which are of unknown real world merit unless that particular simulator's metrics have been demonstrated to bestow significant benefit in the OR.[10,38,39] While a simulator with construct validity can measure performance differences between novices and experts, construct validity represents only a step towards demonstrating a simulator's clinical worth.

The last choice follows the path of moderation. Endovascular procedures, like all surgical endeavors, are not merely psychomotor skills isolated from subjective judgments. Rather, they are a marriage of the two. Thus, intervention is the application of physical skills utilized in an intelligent manner to produce therapeutic results in unpredictable clinical situations. Synergizing subjectivity and objectivity results in a reliable form of skill evaluations. Similar logic led to the development of Objective Structured Assessment of Technical Skills (OSATS), which give the evaluating proctor a validated tool to use when assessing surgical trainees.[40-44]

Skills Transfer

Do skills learned in the virtual environment transfer to the catheterization laboratory?[45-47] The answer to this question is the gold standard which the VR-Lab must meet in order to attain widespread scientific acceptance.[13,48] Virtual reality simulators from other disciplines have shown some predictive value, but those instances are few according to experts.[49,50] Furthermore, although assumed, no specific evidence exists for endovascular simulators demonstrating virtual skills transferability.[50,51] Early attempts to demonstrate transferability of VR skills in other surgical disciplines have often lacked comparable training in the control group, yet have not resulted in definitive evidence as recently pointed out by Schijven et al.[12,39,52,53]

Hence, acceptance of the VR-lab as a venue for skills acquisition within the scientific community awaits evidence demonstrating skills transfer from VR to OR.[32,45-47,50] As rational as this requirement is, the same demands were never placed upon the P-Lab. Remarkably, no evidence supporting or refuting endovascular skills transfer from the research animal model to the operating room (OR) or Cath-Lab exists at the time of this writing. The lack of evidence demonstrating the transferability of skills learned in the P-Lab to the OR leads one to conclude that the gold standard—for non-clinical endovascular interventional training—has never been thoroughly tested in a manner befitting the scientific method.

Cost Effectiveness

The bottom line of the yearly budget has an educational impact on all institutions of higher learning. The demands of training increasingly complex surgical procedures have risen as the available amount of legal working hours has diminished.[54-58] Given that two alternative methods produce similar clinical results, institutions should prefer the economical method over the more expensive one. The ethical disbursement of economical resources—defined here as the highest pragmatic gain from the least financial burden—are a necessity in modern healthcare. We must attempt to provide the best healthcare to the greatest number of individuals using the finite quantities of public or private capital.

Ethics

Since their introduction in 1959, scientists have been striving to adhere to the three "R's" of replacement, reduction and refinement wherever the use of research animals are required by experimental design.[59] In their groundbreaking book, Russell and Burch urged scientists to replace animals with insentient material, i.e. in vitro, or to substitute to a lower species. Furthermore, they plead for a reduction in the number of animals used to obtain the necessary information to the lowest level deemed statistically adequate. Finally, study methodology was to be refined in such a way as to decrease the incidence and severity of pain and distress in any animals that were to be used. Indeed, progress has been made thanks to advances in biomedical cell biology allowing increase use of cell cultures, improved statistical power calculation abilities and increased ethical awareness by researchers when designing studies.

Since then, animal rights activists have been successful in lobbying for more restrictive laws with help from a sympathizing media and an empathizing public. The changes and restrictions brought about by their success are not necessarily deleterious to research. The widespread use of animal care and ethical review boards are prime examples of such progress which justly requires, at the very least, reflection about what is to be gained scientifically from a proposed experiment and how it is to be accomplished.[60] Nevertheless, the experimental use of animals is—despite increased public awareness—frequently accepted as scientific dogma and a necessary evil.[61] Most people generally accept the pragmatic view of mankind as the most valuable species of all. DeGrazia successfully placed both sides of the animal rights argument on common ground when he listed a collection of points on which both sides might agree. He stated that the use of sentient animals, i.e. those capable of experiencing pain and distress, for medical research raises ethical issues and that these animals deserve special care.[62] Regan presented a similar argument when he stated that an individual animal's inherent value does not disappear merely because researchers fail to find an alternative experimental design.[63] In essence, we ought to be thinking of better ways to get the experimental data that humanity needs while including the earlier mentioned three R's at each step.

Many universities provide endovascular training courses using anesthetized research animals to help aspirants over the steepest part of the learning curve-- the window of (in-) opportunity where most surgical errors occur. The basis for research animal training is to maximize patient safety until the novice gains a rudimentary understanding of interventional skills. While curricula vary, a normal course includes didactics followed by hands-on practice in the research animal catheterization laboratory. The adult swine vascular tree approximates human vessels despite anatomical incongruence and lack of pathological disease found in patients. This method offers the advantages of *in vivo* training and the use of the exact equipment with which expertise are being sought. Additionally, encountered complications serve as an enriching experience for trainees as they learn to deal with the unexpected. Virtual reality offers *in vitro* training using slightly modified equipment to treat computer simulated patients with common human pathologies, e.g. carotid arterial stenosis.

SIMULATOR VALIDITY

A critical element of any educational measurement tool, is to assure its validity. The most rudimentary form of validity entails face validity. Face validity is the degree of realism the virtual simulation can mimic. Normally, subject matter experts in the specialty of interest are allowed to subjectively evaluate how well a simulated scenario compares to their real clinical experiences.

An obligate part of confirming a simulator's validity is to establish from a psychometric perspective its construct validity; defined as the ability of a tool to measure the trait it purports to measure. Construct validity is often affirmed and inferred, by establishing that performance improves with experience. For example, the Minimally Invasive Surgical Trainer-Virtual Reality (MIST-VR™) construct validity was confirmed in a study which demonstrated its ability to stratify laparoscopic VR performance based upon individual clinical experience (n=41) using a common procedure, but nevertheless a demanding one.[64] Additionally, many modern endoscopic simulators have gone through generations of improvements and scientific evaluations to make them valuable tools for shortening the learning curve.

A randomized, double blind study conducted by Grantcharov and coworkers using the MIST-VR™ showed decreased surgical errors (p = 0.003) and operative times (p = 0.021) during laparoscopic cholecystectomies for VR trained residents when compared to the control group (n=16).[39] A previous randomized, double blind study done by Seymour et al demonstrated that VR trained surgeons completed laparoscopic cholecystectomies 29% faster (p = 0.039) while non-VR trained surgeons were five times more likely to make errors (p = 0.039, n=16).[65] Furthermore, two separate studies recorded shorter learning curves in the acquisition of the basic psychomotor skills necessary for laparoscopic surgery when training included VR simulators.[20,66]

A high degree of correlation between clinical skills and virtual reality skills is the definition of concurrent validity. A strong correlation between the two measurements makes non-clinical skills assessment and improvement possible in the VR-Lab. At present, there is no generally accepted measurement standard of endovascular interventional skills in the human Cath-Lab skills. Therefore, establishing true concurrent validity is impossible for the moment because there exists nothing to compare new clinical scales against.

Clearly, the advantages and possibilities of VR training have been proven in other surgically focused fields, but, as of yet, few studies have been focused on interventional radiology (IR). The need exists to investigate if current VR technology is capable of assessing the psychomotor skills of the interventionalist as has been done in laparoscopic surgery.[31] If such assessment is possible and the construct validity is verified, then progress towards establishing trainee benchmarks can be begun.[67]

“ GOOD JUDGEMENT COMES FROM EXPERIENCE.

EXPERIENCE COMES FROM BAD JUDGEMENT.”

-UNKNOWN

CATCH 22

Novice interventionalists are thus trapped in a classic Catch 22 proposition; to become proficient in the shortest amount of time at a reasonable price for their respective educational institutions while creating minimal ethical dissonance both in clinical and non-clinical situations. Sound critique, either subjective or objective, and guidance from clinical mentors speeds this journey, yet only independent time in the catheterization laboratory will lead to the vast experience one needs to truly achieve excellence. Table 1 represents the possible venues for this training and experience. Each has its own merits and possibilities, yet many questions remain.

	CATH-LAB	P-LAB	VR-LAB
Psychomotor skills training	✓	✓	✓
Procedural skills training	✓	✓	✓
Cognitive skills training	✓	✓	?
Objective feedback (OSATS or Metrics)	✓	✓	✓
Subjective feedback	✓	✓	✓
Experience with human pathology	✓	×	✓
Patients spared initial learning curve	×	✓	✓
Flexible learning time removed from the clinics	×	✓	✓
Reduced radiation exposure	×	×	✓
Systematic procedural exposure	×	✓	✓
Cost effective	?	?	?
Ethical Concern	?	?	×
Shorter learning curves and reduced error rates	?	?	?
Maximized interventional skill training	?	?	?

✓ = Feasible, × = Infeasible, ? = Unknown

Table 1 Comparison for the alternative endovascular interventional training methods; Catheterization- (Cath-Lab), Porcine- (P-Lab) and Virtual Reality catheterization laboratory (VR-Lab).

AIMS

I. To assess the construct validity of the renal artery stenosis modules of the Procedicus-VIST™ simulator. A) Can the VR simulator stratify interventional performances based upon prior endovascular experience? B) Will the system work as a performance assessment tool outside the real catheterization laboratory? C) Is the VR-Lab useful as a pedagogic tool?

II. To compare the training effects of using the porcine model to virtual reality training in the endovascular novice. A) Is virtual reality simulation as effective a training tool as the porcine laboratory? B) Do skills learned in virtual reality transfer to the catheterization laboratory? C) Is one form of non-clinical training subjectively preferred over the other by trainees?

III. To conduct an economic analysis of the two non-clinical training forms for use in basic endovascular skills training. A) How does the VR-Lab purchase compare financially to the renting of the P-Lab? B) How does the rental of both laboratories affect the relative cost ratios?

IV. To assess the construct validity of the carotid artery modules of the Procedicus-VIST™. A) Can objective metric data from these modules stratify virtual performances based upon experience level? B) Can these simulator modules be used to assess endovascular skills outside of the catheterization laboratory? C) Is the VR-Lab useful as an educational tool?

METHODS AND MATERIAL

SUBJECTS

For study I, eight interventional radiologists and eight medical students undergoing their surgical clerkships participated. The expert group had a mean age of 48 years (range 42-63) and consisted of seven males and one female. Their mean IR experience was 10 years (range 8 months - 20 years). Their mean number of renal artery stenosis (RAS) interventions per year was 11 procedures per year (range 1-40) and two had prior simulator experience. Four played video games “sometimes” and four played “never” on a scale that included: often, sometimes and never. The novice group had a mean age of 28 years (range 24-32) and also consisted of seven males and one female. No one had interventional or previous simulator experience. Six played video games “sometimes” and two played “never.”

Study II enlisted a group of twelve vascular surgeons and interventional radiologists to participate in a two day experimental endovascular training course consisting of the porcine model and VR simulations. The trainees (11 males, one female, 27-61 years old) had a mean open surgical experience of 8 years (range 0-31) and mean endovascular experience of one year (range 0-5 years). Two participants had limited prior virtual reality simulation exposure (<15 minutes). A group of six experienced interventional radiologists with a mean interventional experience of 10.5 years were recruited to function as onsite proctors. Each proctor evaluated the same trainees in the VR-Lab and the P-Lab to minimize variability. Two highly experienced interventional radiologists with a mean of 22.5 years of interventional experience were recruited to serve as the video assessor panel.

Figure 1 Procedicus-VIST, Mentice Medical Simulations, Gothenburg, Sweden

1. Insert a 0.035” J-profile guide-wire and a 4/5 F Pigtail diagnostic catheter over the 0.035” guide-wire into the distal aorta.
2. Connect the contrast line and perform a Digital Subtraction Angiography (DSA) of the iliac arteries via the Pigtail.
3. Use the 0.035” guidewire and Pigtail to canalize the contralateral common iliac artery. (You may need to change catheters, e.g. Sim1 or SHK1, or guidewires, e.g. hydrophilic or stiff, in order to accomplish this.)
4. Insert a 6/7 F guiding catheter into the contralateral common iliac artery proximally to the stenosis.
5. Carefully transverse the “stenosis” with the guidewire.
6. With the guidewire still in position, connect the contrast line, perform a selective DSA or roadmap of the contralateral external iliac artery.
7. Measure and evaluate the external iliac artery and the “stenosis.”
8. Insert an appropriately sized peripheral stent catheter over the wire.
9. Center the stent within the lesion and carefully deploy the stent.
10. Maintain your distal guidewire position and remove the stent catheter.
11. Insert an appropriately sized peripheral dilation balloon catheter over the 0.035” guidewire, advance into the stent’s lumen and perform a post-deployment PTA.
12. Maintain the distal guidewire position and remove the angioplasty catheter.
13. Connect the contrast line and perform a control DSA via the guiding catheter.
14. Withdraw the 0.035” wire and introducer.

Table 2 Iliac artery stenting protocol.

Study III was conducted using economic data extracted from Study II but included no new participants.

Study IV comprised experienced interventionalists and medical students during their surgical clerkship. The expert group had a mean age of 49 years (range 36-65) and consisted of fifteen males and one female. Their mean IR experience was 11 years (range 1-25). The novice group had a mean age of 29 years (range 23-39) and consisted of thirteen males and three females. Five had limited previous simulator experience yet none had any IR experience. Neither group had any experience in placing carotid artery stents.

1. Place .035” J-type guidewire & .035” Pigtail catheter into the ascending aorta.
2. Position C-arm in LAO (25°-60°), remove .035” guidewire and perform DSA with AP and lateral views.
3. Replace .035” wire, exchange to .035” diagnostic catheter and position the guidewire in the external carotid artery.
4. Insert a 6F guide catheter proximally to the CCA bifurcation and remove the .035” guidewire.
5. Insert the .014” EPD guidewire, transverse the stenosis and deploy the EPD filter.
6. Using a .014” peripheral balloon catheter pre-dilate the lesion.
7. Insert .014” carotid stent catheter, deploy the stent and perform a DSA.
8. Repeat PTA if necessary.
9. Use a .014” recovery sheath catheter to collect the EPD.
10. Perform control DSA of the right ICA including intracranial views.
11. Check for spasm? Dissection? Remove all equipment.

Table 3 Carotid artery stenting protocol for the right internal carotid artery using a femoral artery approach.

VIRTUAL REALITY LABORATORY

The virtual reality simulator used in all experiments was the Procedicus-VIST™ system (Mentice Medical Simulations, Gothenburg, Sweden) which consisted of a double processor computer, a touch-sensitive screen, a viewing screen and a simulator dummy as seen in **Figure 1**. The dummy concealed various mechanical systems, which registered the physical movements of and produced the haptic feedback to the users. The touch screen was driven by a menu system which allowed the trainees to select guide wire, diagnostic catheter, guiding catheter or stent/balloon catheter by type and diameter. Fluoroscopic view, instrument selection, total IV contrast dose and intervention time were continuously displayed on the viewing screen. The simulator dummy consisted of a plastic human form in the supine position with a right femoral artery port lying upon a catheterization laboratory type table. An introducer was permanently placed within the right femoral artery. A standard two-pedal system with X-Ray and Cine lay under the table. A double joystick control box allowed the operator to; 1) Position the virtual fluoroscope 2) Zoom the view 3) Replay cine sequences 4) Capture and simultaneously display an overlapping roadmap on the viewing screen 5) Reposition the catheterization table virtually. An inflator apparatus with a pressure gauge was permanently attached to the main system. When either a balloon or stent catheter was selected and introduced into the dummy, the inflator produced appropriate morphological changes on screen. Once deployed, stents remained in place for the entire sequence. A permanently attached IV contrast syringe prepared with a one-way valve was used for simulated contrast infusion. Replenishment was accomplished simply by retracting the plunger, which drew air in place of real contrast into the syringe. Real catheters and guidewires in sizes ranging from 0.014”-0.035” (5-7 French (Fr)) were used. The profile/flexible ends could not be inserted into the machine due to the simulator’s mechanical design. Instead, straight ends were used in order to properly engage the haptic mechanisms within the device. The program automatically displayed the correct tip profile for the selected instrument on the viewing screen.

PORCINE LABORATORY

In three separate catheterization laboratories, normally fed Swedish swine were sedated with i.m. ketamine and dormicum and ventilated on a mixture of oxygen and N_2O. Anesthesia was maintained with continuous i.v. infusion of thiopenthatol and buprenorphine supplemented with isoflurane as necessary. After catheterization of the right femoral artery using the Seldinger technique and a 10Fr introducer sheath (Cook Inc, USA), an i.v. bolus dose of 300U/kg of Heparin was given. Repeat anticoagulation was given every 60 minutes thereafter. Preparation for iliac artery stenting was accomplished using various equipment from Cordis, Johnson & Johnson, USA and Boston Scientific, USA. The final stenting and angioplasty was performed with a nitinol stent and an optiplast balloon from Bard, Inc, USA (Luminex® and XT Optiplast®). Fluoroscopy and Digital Subtraction Angiography (DSA) was performed using Philips and GE fluoroscopic equipment (Philips Medical Systems, Eindhoven, The Netherlands and General Electric Healthcare, UK). Using sodium ioxaglat (Omnipaque® 240 ml) as the contrast medium, the size of the left external iliac artery and the “stenosis” to be treated was estimated using a semi-quantitative method where the guiding catheter was used as a reference. The “stenosis” was delineated by the proctor on the fluoroscopic screen using transparent, self-adhesive plastic book marking tabs. At the conclusion of the experiment the pigs were euthanized with a lethal IV bolus of potassium chloride.

LOGISTICS

Studies I & IV

Prior to construct validations, all participants received a 45 minute, standardized didactic introduction to the simulator and the requisite endovascular techniques. Participants were given the same modules to complete, the rules for completion and the role of the proctors. The objective of the procedure was to revascularize the afflicted artery using either balloon angioplasty, stents or a combination of both. Study I used six different renal artery stenosis (RAS) modules, i.e. simulated patients, which were completed twice without the aid of embolic protection devices, once during familiarization and again during the testing phase. Study IV assessed performance using a unique carotid artery stenosis (CAS) module, i.e. it was not available during familiarization training, with the aid of an embolic protection device. Any case which took over 30 minutes (I) or 60 minutes (IV) to complete was abandoned, during familiarization and testing, to allow attempts at the remaining cases. A short journal was presented onscreen and the intervention timer automatically started once the operator began instrument selection. The operators were requested not to be overly concerned with speed and to complete each procedure to the best of their ability. All subjects attempted each procedure twice, once during the familiarization period and again during an undisturbed test period. In order to mitigate knowledge-based bias, full disclosure of the performance metrics to be recorded was given to both groups, as it was assumed that experts, but not novices, might have had prior knowledge of important metrics within interventional radiology procedures.

Study II

Trainees were informed that they would be participating in an experiment involving proctored evaluations and randomized training methods. Prior to arrival, each received an information package containing the iliac artery stenosis (IAS) procedure protocol, instructions regarding their responsibilities during the experiment and a demographic questionnaire used to experience-stratify and randomize the trainees into four alternative training groups. Proctors and video assessors received a one hour introduction to the trainee evaluation methods prior to the study's commencement. All participants received a one hour didactic introduction to the simulator and the necessary endovascular techniques before the evaluations began. The interventional objective was to revascularize the iliac artery with balloon angioplasty and nitinol stents using the standardized protocol, **Table 2**. Any case which took over 30 minutes to complete was aborted. Following initial evaluations, training groups completed two unevaluated training sessions of three hours each according to their randomized method(s), **Figure 2**.

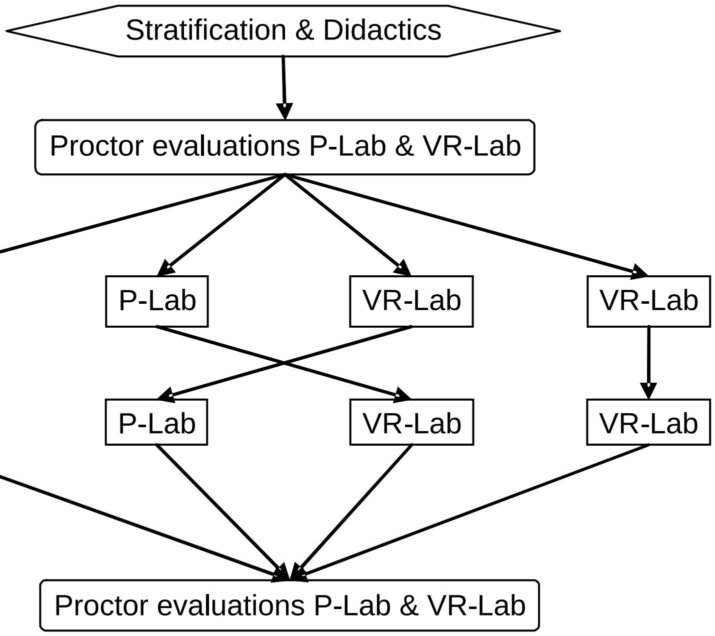

Figure 2 Study II, training flowchart.

Study III

Actual cost data were collected from the university P-Lab and VR-Lab during the two day course conducted in October 2005. Rental fees for both labs included full ancillary support for the swine and simulators. Proctors and video assessor salaries were not included in this analysis as teaching was assumed to be a regular function of an academic institution. Each proctor attended both days of the course, totaling 16 hours per course, while the video assessors took one eight hour day to complete video recording reviews. Market research and formal manufacturer quotes were used to complete any missing prices. The costs of the VR-Lab was chosen as the numerator and the P-Lab as the denominator for the comparison ratio. A five year analysis was chosen as the longitudinal length of comparison to reflect the expected life cycle of the simulator.[68] The original value of the Procedicus-VIST™ was assumed to be €200,000. Residual value was presumed to 25% of the original value, i.e.

€50,000, after the five year life cycle ended making the depreciation value equal to €150,000 with an annual depreciation value of €30,000 using straight-line depreciation. Laboratory rental prices were assumed to increase at a rate of 2.5% per year to match the current Euro inflation rate. Annual national IR and vascular surgical training demand was calculated as 52 and estimated to increase at a 2% per year. Sensitivity analysis— an essential tool used in making economical decisions to clarify the impact of price variability—was performed to assess the impact of price variations for each laboratory given a 50% rise or fall in costs. Such an analysis helps safeguard against making incorrect decisions based upon singularly interpreted financial data. The first year annual difference was assumed as the short-term potential savings and the five year cumulative total as the long-term potential savings.

ENDPOINTS

Objective Endpoints

Objective performance metrics for studies I and IV were automatically recorded for each case during the testing phase only by the Procedicus-VIST™ software, **Table 4**. The Total Score performance evaluation forms used in study II consisted of modified surgical evaluation forms originally developed for surgical skills assessment.[38]. The iliac artery procedure forms used were the Task Specific Checklist (TSC) and the Global Rating Scale (GRS), which yielded a Total score. The TSC was completed during the procedure, **Table 5**. Correct completion of a task resulted in the assignment of one point (max=14). If the task was improperly performed or omitted, the trainee received zero points. Additionally, if the trainee asked for help or the proctor needed to intervene to prevent a catastrophic event, the trainee received a zero for that task, even if it was completed properly afterwards. The GRS, based on multiple questions and a five point anchored Lichert scale, was filled out once the procedure was completed to ensure an overview of the entire performance sequence (max=45), **Table 6**. Proctors performed the evaluations during and immediately after the procedures were completed. Video assessors evaluations, performed using blinded video recordings from the VR-Lab, were used to calculate Total Score, Task Specific Checklist and Global Rating Scale inter-rater reliability. Blinded video ensured that each participant's identity was protected during videography.

Procedure Time	Minutes needed to complete the entire procedure
Fluoroscope Time	Minutes the fluoroscope was used during the procedure
Contrast	Total amount of contrast medium used, in milliliters
Cine Loops	Number recorded during procedure
Lesion Coverage %	Percentage of lesion covered by selected tool
Tool : Lesion Ratio	Inflated tool's diameter : lesion's diameter
Placement Accuracy	Distance, in millimeters, from the actual placement of the tool to the lesion's center, longitudinally
Residual Stenosis	Percentage stenosis post PTA or stent deployment

Table 4 Metric definitions.

	OMITTED OR INCORRECT	DONE CORRECTLY	INSTRUCTION REQUIRED*
1. Positioned guidewire and diagnostic catheter correctly?	0	1	
2. Distal aorta DSA conducted correctly?	0	1	
3. Proper guidewire technique used to gain contralateral access?	0	1	
4. Guiding catheter properly placed in the contralateral common iliac artery?	0	1	
5. Transversed "stenosis" with correct technique?	0	1	
6. External iliac artery DSA / roadmap conducted properly?	0	1	
7. Measurements and evaluations performed accurately?	0	1	
8. Proper stent catheter selection?	0	1	
9. Deployed stent accurately?	0	1	
10. Maintained guidewire position across lesion?	0	1	
11. PTA conducted correctly?	0	1	
12. Maintained guidewire position across lesion?	0	1	
13. Control DSA conducted correctly?	0	1	
14. Extraction of guidewire and guiding catheter performed correctly?	0	1	

Table 5 Iliac artery stenosis task specific checklist. * Receiving instruction resulted in a zero even if the step was correctly executed afterward.

Aseptic Technique*				
1	2	3	4	5
Sloppy with high risk for contamination		Reasonable but some lapses hat risk loss of sterility		Careful with little risk of compromising sterility
Respect for Tissue				
1	2	3	4	5
Frequently uses unnecessary force or causes damage		Careful handling but occasionally causes inadvertent damage		Consistently handles with minimal damage
Fluoroscopic proficiency				
1	2	3	4	5
Poor control and selection of inappropriate view or causes patient injury		Competent use but with some lapses in control or sub-optimal views		Prompt attainment of appropriate fluoroscopic views
Time & Motion				
1	2	3	4	5
Slow with many unnecessary moves and instrument changes		Makes reasonable progress but some unnecessary moves		Clear economy of movement and maximum efficiency
Instrument Handling & Safety				
1	2	3	4	5
Repeatedly makes tentative, awkward or unsafe moves		Competent use but occasionally awkward or tentative		Fluid movements without stiffness
Knowledge of Instruments*				
1	2	3	4	5
Frequently asks for or uses wrong instrument		Knows names of most instruments and uses them properly		Obviously familiar with all instruments and their uses
Knowledge of Specific Procedure*				
1	2	3	4	5
Require specific instruction for most steps		Knows all the important steps		Demonstrates familiarity with all steps
Quality of Final Product				
1	2	3	4	5
Well below standard and likely to fail		Deficiencies but would probably function adequately		Excellent with no flaws and likely to function well
Radiation Discipline				
1	2	3	4	5
Patient and Operator repeatedly overexposed. Poor use of shutters and shielding.		Aware of exposure but needs improvement. Adequate shutter and shield use		Judicious use of fluoroscopy. Exposure kept to a bare minimum.

Table 6 Iliac artery stenosis global rating scale. * Not possible to evaluate on video.

Subjective Endpoints

Participants completed exit surveys to record their demographics and subjective opinions regarding the simulator and training formats. For study I, the questionnaire required the participants to use a visual analogue scale (VAS) to rate the following eight parameters: 1) Total realism 2) Guidewire realism 3) Catheter realism 4) Balloon realism 5) Stent realism 6) Fluoroscopic realism 7) Joystick control realism and 8) Pedagogic instrument. VAS scores marked on a 10 centimeter line ranged from zero to 100, where a higher score meant a better rating. For studies II and IV, trainees completed exit surveys, again using the VAS system, ranging from zero (strongly disagree) to 100 (strongly agree) to indicate their agreement to statements regarding the two training methods, **Table 7**.

VR training is better than the porcine lab
I understand the IAS procedure better than I did before the course
VR training could replace the porcine lab
VR simulation training should be obligatory
Porcine lab training should be obligatory
I could have completed the procedure equally well without the Proctor
VR is a valuable training instrument
The VR simulations were realistic compared to the porcine lab
Porcine lab training is better than the VR lab

Table 7 Exit survey statements.

STATISTICAL ANALYSIS

Data analysis was performed using the statistical program "R", an open-sourced language statistical computing and graphics program.[69] For studies I and IV, the difference of interest was defined as 2 SD from the expert mean for each objective parameter. The large difference was used to detect a presumably vast performance difference between the groups based upon their endovascular experience levels and to maintain a small study size. Only results from the testing period were analyzed. No analysis for correlation for video gaming was performed. Student's t-test was used for analysis of the metric means. Study II included an analysis of covariance (ANCOVA) to assess the effect of the independent variables-- previous number of P-Lab/VR-Lab sessions, interventional experience and surgical experience-- on the dependent variable Total-score. The dependence of the multiple observations for each person, i.e. clustered data, was accounted for through the use of generalized estimating equations (GEE).[70] GEE provided adjusted standard errors for evaluating significance of terms in the ANCOVA models. Video assessor evaluations from the VR-Lab were compared to the proctor evaluations to establish inter rater reliability (IRR).[42,71] Statistical significance was assumed to be present when $p < 0.05$. Study III 's economical analysis was conducted according to the National Institute of Health's recommendations for the implementation of information technology (IT) products.[68] Economical sensitivity analysis consisted of increasing or decreasing the estimated costs for each laboratory by 50% to produce varying cost ratios.

ETHICAL APPROVAL

As studies I, III and IV were conducted outside of the research and human catheterization laboratories, no ethical review board decision was sought. Study II was approved by the local ethics review board and conformed to the *Guide for the Care and Use of Laboratory Animals* published by the U.S. National Institute of Health (NIH).[72]

RESULTS

PAPER I

Table 8 contains the objective metric performance data for each parameter by group for each of the six cases. Statistical analysis across the different modules revealed no significant differences in performances between the two groups in procedure time, number of cine loops, lesion coverage, tool to lesion ratio, placement accuracy and residual stenosis. The total fluoroscopic use was greater for the novice group ($p < 0.01$). One expert failed to finish a module within the 30 minute cutoff limit. All others participants completed the assigned modules on time.

		CASE 1	**CASE 2**	**CASE 3**	**CASE 4**	**CASE 5**	**CASE 6**
Procedure Time (mins)	**E**	**8.4 (2.1)**	**9.6 (4.0)**	**13.1 (4.1)**	**11.4 (2.6)**	**7.9 (1.8)**	**10.1 (2.2)**
	N	**12.7 (4.6)**	**16.3 (3.7)**	**14.9 (5.9)**	**13.9 (4.0)**	**11.0 (3.8)**	**11.9 (4.0)**
	P	**<0.02**	**<0.01**	**NS**	**NS**	**<0.05**	**NS**
Fluoro time (mins)	**E**	**3.5 (1.3)**	**4.4 (1.5)**	**5.8 (1.4)**	**4.9 (2.0)**	**3.0 (0.7)**	**4.7 (1.3)**
	N	**7.7 (2.9)**	**9.5 (3.1)**	**9.1 (4.8)**	**9.8 (4.2)**	**6.4 (2.1)**	**7.5 (2.7)**
	P	**<0.001**	**<0.001**	**<0.05**	**<0.01**	**<0.01**	**<0.02**
Cine loops (n)	**E**	**3.1 (1.1)**	**3.6 (1.6)**	**6.3 (3.2)**	**3.6 (1.3)**	**3.3 (1.7)**	**4.9 (2.4)**
	N	**3.6 (2.6)**	**4.8 (3.9)**	**3.9 (2.0)**	**3.0 (2.2)**	**3.8 (2.8)**	**2.9 (1.6)**
	P	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**
Lesion coverage (%)	**E**	**99.2 (1.9)**	**94.8 (5.8)**	**98.0 (5.2)**	**83.1 (12.8)**	**90.9 (12.5)**	**58.8 (32.9)**
	N	**97.4 (3.8)**	**93.4 (8.7)**	**94.1 (7.4)**	**85.6 (9.5)**	**95.8 (6.9)**	**68.5 (22.9)**
	P	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**
Tool : vessel ratio (%)	**E**	**100.0 (9.0)**	**90.0 (18.2)**	**109.2 (16.3)**	**99.6 (16.3)**	**87.9 (8.2)**	**98.3 (13.2)**
	N	**90.3 (24.1)**	**87.1 (26.9)**	**100.3 (23.4)**	**91.4 (16.3)**	**84.4 (11.7)**	**84.1 (14.7)**
	P	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**
Placement accuracy (mm)	**E**	**1.0 (1.2)**	**2.3 (1.8)**	**1.2 (1.2)**	**3.0 (1.6)**	**1.9 (1.2)**	**6.8 (5.3)**
	N	**1.5 (0.7)**	**2.1 (1.7)**	**1.8 (1.1)**	**2.9 (0.8)**	**1.3 (1.3)**	**4.6 (3.6)**
	P	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**
Residual stenosis (%)	**E**	**2.8 (4.3)**	**14.1 (12.8)**	**4.7 (6.4)**	**6.7 (8.9)**	**12.9 (6.2)**	**5.8 (8.6)**
	N	**13.6 (20.7)**	**17.7 (21.5)**	**11.1 (14.5)**	**11.9 (12.9)**	**15.7 (11.4)**	**16.1 (14.5)**
	P	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**

Table 8 Study I, means and standard deviations (SD) for experts (E) and novices (N) with p-values (P). NS = not significant.

Table 9 contains the subjective rating data from the exit survey. The mean VAS scores rating the use of the VIST™ as a pedagogic instrument was 89 (SD 9) for the experts and 92 (SD 9) for the novices. Statistical analysis revealed no significant differences in mean VAS scores for total realism, guidewire realism, catheter realism, balloon realism, stent realism, fluoroscopic realism and use as a pedagogic instrument. Mean VAS scores for joystick realism was rated significantly higher by the expert group ($p < 0.02$).

	EXPERTS	NOVICES	P-VALUE
Total realism	61 (27)	68 (16)	NS
Guidewire realism	71 (29)	85 (9)	NS
Catheter realism	70 (30)	83 (13)	NS
Balloon realism	75 (24)	82 (12)	NS
Stent realism	59 (34)	77 (12)	NS
Fluoroscopic realism	95 (9)	86 (10)	NS
Joystick realism	94 (7)	71 (24)	< 0.02
Pedagogic instrument	91 (9)	92 (9)	NS

Table 9 Study I, VAS scores (0-100) with means and standard deviations (SD).

PAPER II

Total Score IRR, determined by calculating the inter-rater correlation coefficient (ICC) between the proctors and video assessors, was found to be substantial at a value of 0.679. Cumulative VR-Lab sessions improved VR-Lab Total scores ($\beta = 3.0$, $p = 0.0015$) and P-Lab Total scores ($\beta = 1.8$, $p = 0.0452$), **Table 10**. P-Lab sessions improved P-Lab Total scores ($\beta = 4.1$, $p < 0.0001$) but had no effect on VR-Lab Total scores. In the general statistical model for Total scores from all laboratories, both P-Lab sessions ($\beta = 2.6$, $p = 0.0010$) and VR-Lab sessions ($\beta = 2.4$, $p = 0.0032$) significantly improved Total scores. Neither previous surgical nor IR experience affected Total scores. VR-Lab Total scores were consistently higher than P-Lab scores ($\Delta = 6.7$, $p < 0.0001$). Analysis of mean Total scores in each laboratory failed to detect any ceiling effect, **Figure 3**.

Total Score	**B**	**95% CI**	**P-value**
Surgical Experience	- 0.2	- 0.5, 0.1	0.1141
IR Experience	1.4	- 0.8, 3.6	0.2295
P-Lab sessions	2.6	1.0, 4.1	0.0010
VR-Lab sessions	2.4	0.8, 4.1	0.0032
VR-Lab Total Scores			
Surgical Experience	- 0.1	- 0.4, 0.1	0.2612
IR Experience	1.6	- 0.3, 3.5	0.0921
P-Lab sessions	1.3	- 0.2, 2.8	0.0902
VR-Lab sessions	3.0	1.1, 4.9	0.0015
P-Lab Total Score			
Surgical Experience	- 0.3	- 0.7, 0.0	0.0826
IR Experience	1.2	-1.5, 3.9	0.3923
P-Lab sessions	4.1	2.1, 6.1	< 0.0001
VR-Lab sessions	1.8	0.0, 3.9	0.0452

Table 10 Study II, Total score ANCOVA partial coefficients.

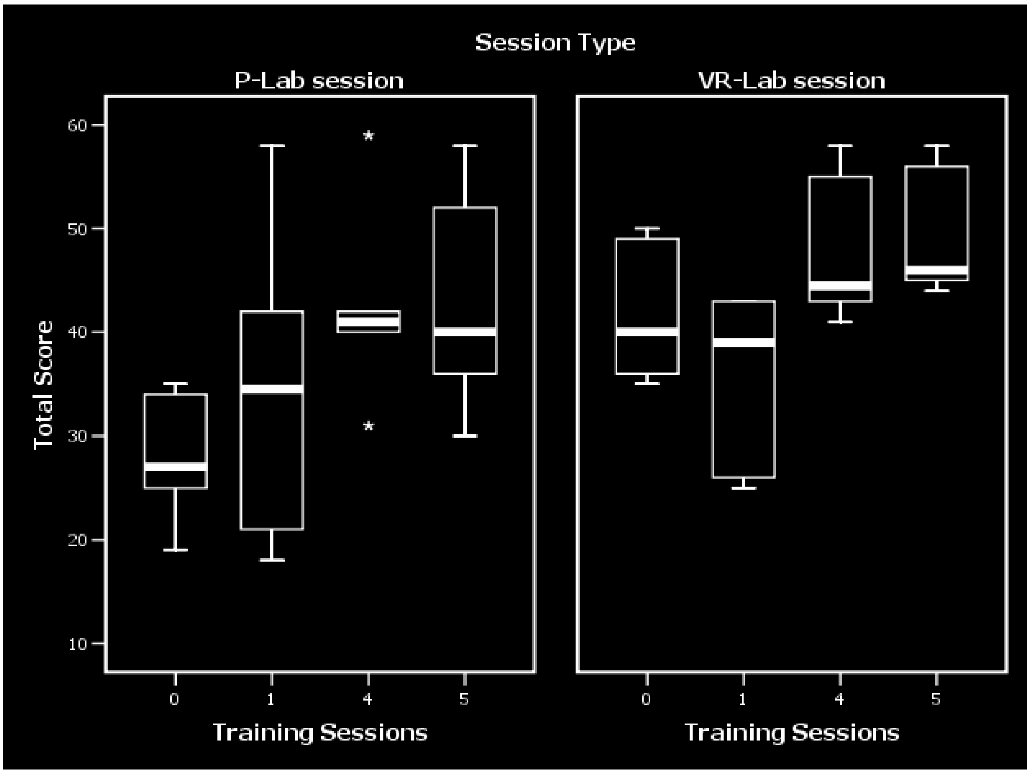

Figure 3 Study II, Total score box-whisker plots.

Ten aborts-- seven failures to finish (58%) and 3 dissections (25%)-- occurred during the initial evaluations in the P-Lab (completion rate=17%). Five aborts, all failures to finish (42%), occurred during the VR-Lab initial evaluations (completion rate=47%). Two aborts-- one perforation (8%) and one failure to finish (8%)-- occurred during the P-Lab final evaluations (completion rate=83%). All trainees completed the IAS procedure within the allotted time in the final evaluations in the VR-Lab.

Subjects agreed more with the statement that P-Lab training was better compared to the VR-Lab training (mean 87, SD 17) than to the reverse statement (mean 17, SD 18), **Table 11**. The majority agreed that both forms of training should be obligatory; P-Lab (mean 79, SD 31), VR-Lab (mean 79, SD 17). Realism ratings for the VR-Lab compared to the P-Lab were low (mean 37, SD 23) and most disagreed with the statement that VR-Lab training was capable of replacing the P-Lab (mean 15, SD 17). However, most agreed that the VR was a valuable training instrument (mean 80, SD 25). Few agreed with the statement that they would have been able to complete the IAS procedure equally well without a proctor (mean 17, SD 29). Finally, most agreed that they understood the procedure better after the course than they did upon arrival (mean 81, SD 30).

	MEAN (SD)
VR training is better than the porcine lab	17 (18)
I understand the IAS procedure better than I did before the course	81 (30)
VR training could replace the porcine lab	15 (17)
VR simulation training should be obligatory	79 (17)
Porcine lab training should be obligatory	79 (31)
I could have completed the procedure equally well without the Proctor	17 (29)
VR is a valuable training instrument	80 (25)
The VR simulations were realistic compared to the porcine lab	37 (23)
Porcine lab training is better than the VR lab	87 (17)

Table 11 Study II, exit survey results . VAS scores (0=strongly disagree, 100=strongly agree).

PAPER III

Using a VR-Lab purchase vs. P-Lab rental analysis, the cost of the VR-Lab for a two day course given to 12 study participants was €34,348 or €2,862 per trainee, Table 12. In contrast, the pig lab cost €46,350 to for these same 12 participants, Table 13. This translates to an approximate cost of €3,862 per trainee. The cost ratio of the VR-Lab to the P-Lab was found to be 0.74 in favor of the VR-Lab, Table 14. Sensitivity analysis resulted in a cost ratio range of 0.25 in favor of the VR-Lab to 2.22 in favor of the P-Lab. The first year potential national savings amounted to €52,009 assuming exclusive use of the VR-Lab for 52 course participants, Table 15. At the end of the five-year, inflation-adjusted period, the cumulative training savings totaled €325,314 excluding a residual value of €50,000 for the purchased VR system.

Using a rental analysis, the cost of the VR-Lab for a two day course was €10,768 or €897 per trainee. The P-Lab costs remained the same. The cost ratio of the VR-Lab to the P-Lab was found to be 0.23 in favor of the VR-Lab. Sensitivity analysis resulted in a cost ratio range of 0.08 to 0.70 in favor of the VR-Lab. The first year potential national savings amounted to €154,189. At the end of the five-year, inflation-adjusted period, the cumulative training savings totaled €844,365.

The largest item cost in the P-lab was the intravascular stents (€38.640, 83% of total) while these constituted a relatively small amount of the VR-Lab budget (€2,140, 6% of total if purchased or 20% using rental analysis). The largest item cost in the VR-Lab was either the annual purchase price (€30,000, 87%) or the rental fees (€6,420, 60%) for the VIST simulator compared to the rental fees for the P-Lab (€4,140, 9%).

Reusable	Number	Price (€)	Total (€)
.035” Guidewire	4	37	148
.014” Guidewire	4	50	200
5 F Diagnostic, short	4	33	132
5 F Diagnostic, long	4	33	132
10 F Introducer	4	36	144
8 F Crossover, 43 mm	4	68	272
8 F Crossover, 33 mm	4	68	272
8 F Guide, 70 mm	4	50	200
Luminex Stent	4	535	2,140
.035” XT PTA balloon	4	77	308
.014” XT PTA balloon	4	100	400
		Subtotal	**4,348**

Purchase	Number	Annual Price (€)	Total (€)
Procedicus-VIST™	1	30,000	**30,000**
		Grand Total	**34,348**
		Individual Cost	**2,862**

Rental	Number	Annual Price (€)	Total (€)
Procedicus-VIST™	3	2,140	**6,420**
		Grand Total	**10,768**
		Individual Cost	**897**

Table 12 Study III, VR-Lab training cost per individual per course. Purchase and rental individual costs listed separately.

Reusable	Number	Price (€)	Total (€)
Puncture Needles	6	6	36
10F Introducer	6	36	216
.035” Guidewire, 160 cm, straight	12	37	444
.035” Guidewire, 260 cm, straight	6	18	108
.035” Guidewire, 260 cm, curved	6	19	114
8F Balkin sheath, “Up&Over”	12	68	816
.035” Terumo, 160 cm, straight	6	27	162
.035” Terumo, 160, curved	6	27	162
.035” Terumo, 260 cm	6	68	408
5F Contra/”Pigtail”	12	33	396
5F Cobra catheter	12	20	240
4F Cobra catheter	6	20	120
5F Sim 1 catheter	6	20	120
5F Bernstein catheter	6	20	120
Pressure Guide Handle	6	18	108
		Subtotal 1	**3,570**

Consumables, Evaluations	Number	Price (€)	Total (€)
Luminex, Nitinol Stent	24	535	12,840
Omnipaque contrast (100 ml)	6	83	498
		Subtotal 2	**13,338**

Consumables, Training	Number	Price (€)	Total (€)
Luminex, Nitinol Stent	36	535	19,260
Saxx Stent	12	385	4,620
Optima XT PTA balloon	12	77	924
Omnipaque contrast (100 ml)	6	83	498
		Subtotal 3	**25,302**

Rental Fees	Number	Price (€)	Total (€)
Catheterization Laboratories	3	440	1,320
Experimental swine	6	413	2,478
Infusion pumps	6	13	78
Tandem Fluoroscopy	6	44	264
		Subtotal 4	**4,140**

Grand Total	**46,350**
Individual Cost	**3,863**

Table 13 Study III, P-Lab training cost per individual per course.

VR-Lab Purchase **P-Lab Rental**	**VR-Lab** **P-Lab**		
	150% VR-Lab	**100% VR-Lab**	**50% VR-Lab**
150% P-Lab	**0.74**	**0.49**	**0.25**
100% P-Lab	**1.11**	**0.74**	**0.37**
50% P-Lab	**2.22**	**1.48**	**0.74**

VR-Lab Rental **P-Lab Rental**	**VR-Lab** **P-Lab**		
	150% VR-Lab	**100% VR-Lab**	**50% VR-Lab**
150% P-Lab	**0.23**	**0.15**	**0.08**
100% P-Lab	**0.35**	**0.23**	**0.12**
50% P-Lab	**0.70**	**0.46**	**0.23**

Table 14 Sensitivity analysis. Cost ratios at various percentages of the cost estimates.

VR-Lab Purchase **P-Lab Rental**	**Lifecycle Course Demand**				
	Year 1	**Year 2**	**Year 3**	**Year 4**	**Year 5**
Course Participants	52	53	54	55	56
Individual Cost :					
VR-Lab	2,862	2,862	2,862	2,862	2,862
P-Lab	3,863	3,959	4,058	4,159	4,263
Annual Cost :					
VR-Lab Total	148,841	151,818	154,855	157,952	161,111
P-Lab Total	200,850	209,989	219,543	229,532	239,976
Annual Difference	**52,009**	**58,171**	**64,689**	**71,581**	**78,865**
5 Year Difference					**325,314**

VR-Lab Rental **P-Lab Rental**	**Course Demand**				
	Year 1	**Year 2**	**Year 3**	**Year 4**	**Year 5**
Individual Cost :					
VR-Lab	897	920	943	966	990
P-Lab	3,863	3,959	4,058	4,159	4,263
Annual Cost :					
P-Lab Total	46,661	48,784	51,004	53,325	55,751
VR-Lab Total	200,850	209,989	219,543	229,532	239,976
Annual Difference	**154,189**	**161,204**	**168,539**	**176,208**	**184,225**
5 Year Difference					**844,365**

Table 15 Five year projected national savings in Euros (€) for interventional radiologists and vascular surgeon course using purchased or rented VR-Lab in place of P-Lab.

Paper IV

Of the seven recorded metrics, only chronologic measurements were found to be significantly better amongst the experts, Table 16. Procedure and fluoroscopic time was 8.7 and 8.7 minutes greater in the novice group, $p = 0.0066$ and $p = 0.0031$ respectively. There were no significant differences in performance between the two groups for the metrics of cine loops, tool:vessel ratio, coverage percentage, placement accuracy or residual stenosis. The differences in the means of these performance metrics on occasion showed a tendency towards a statistically significant difference. For example, experts recorded fewer cine loops, had better residual stenosis and demonstrated better tool:vessel ratios, yet they had worse placement accuracy and poorer lesion coverage. All participants completed the procedure in the allotted time frame.

Contrast medium measurement metrics were found to be too imprecise for comparative statistical analysis. Specifically, volume measurements varied widely for variable injection rates and volumes which resulted in substantially lower cumulative volumes than expected, Table 17.

Table 18 shows the results of the questionnaire administered to participants soliciting their views about VIST™. Students and experts did not differ on four of ten dimensions, including; frequency of video gaming habits, belief that the VIST™ is a good pedagogic instrument, belief that obligatory VR-Lab training is desirable and the desire for proctored training. Amongst the significantly different VAS score categories, novices tended to view VR-Lab training more positively compared to reading alone. They also rated the simulations more realistic, favored them over didactic/AV instruction and left the course with a better understanding of the CAS procedure. However, novices would have preferred to have more time with the simulator while the experts felt adequately familiarized. Novices expressed an increased interest in endovascular medicine whereas the expert group did not.

	EXPERTS	STUDENTS	*P-VALUE*
Total Time (min)	38.6	47.3	0.0066
Fluoro Time (min)	18.0	26.7	0.0031
Cine Loops (n)	11.0	12.4	0.3214
Placement Accuracy (mm)	4.1	2.4	0.1334
Residual Stenosis (%)	22.3	28.7	0.1405
Tool to vessel ratio	77.8	71.3	0.1356
Coverage (%)	90.8	94.8	0.5206

Table 16 Study IV, metric means compared using Student's t-test.

Repeated Injection of 20 ml @				**Repeated Injection of 10 ml @**			
20 ml/sec		**5 ml/sec**		**10 ml/sec**		**5 ml/sec**	
Observed Metric	Expected Metric	Observed Metric	Expected Metric	Observed Metric	Expected Metric	Observed Metric	Expected Metric
4.2	20	12.6	20	2.7	10	6.4	10
8.6	40	25.6	40	5.4	20	12.9	20
12.7	60	37.4	60	8.2	30	19.3	30
17.0	80	49.9	80	10.9	40	25.8	40
21.0	100	62.5	100	13.7	50	32.1	50
25.2	120	75.0	120	16.4	60	38.5	60
29.3	140	87.5	140	19.0	70	45.0	70
33.3	160	100.0	160	21.7	80	51.5	80
37.4	180	112.6	180	24.5	90	58.0	90
Total	Total	Total	Total	Total	Total	Total	Total
41.5	200	125.3	200	27.1	100	64.5	100

Table 17 Study IV, observed versus expected contrast medium volume measurements (ml) at varying boluses and injection rates.

	EXPERTS	**STUDENTS**	**P-VALUE**
I play video games often	8.8	21.0	0.0756
VR training is better than reading	70.5	93.6	0.0004
I now understand CAS better	62.4	89.7	0.0024
VR could replace didactics/AV instruction	52.3	80.1	0.0014
VR training should be obligatory	73.4	83.4	0.1521
I had adequate time with VIST	70.1	44.1	0.0074
I had no need for the facilitator	16.6	30.8	0.1191
The VIST is a good pedagogic instrument	80.4	86.4	0.2848
The simulation was realistic	67.5	82.5	0.0104
I am now more interested in endovascular medicine	38.6	80.3	0.0001

Table 18 Study IV, exit survey data. VAS scores (0 strongly disagree-100 strongly agree).

DISCUSSION

SKILLS ASSESSMENT

The psychometric validation of any virtual reality machine-generated metrics is critical in order to understand the strengths and limitations of emerging educational tools. Studies I and IV attempted to show construct validity and specifically chose two populations, in which we anticipated large differences and hence hypothesized the need for a relatively small sample size. This notwithstanding, essentially no differences were seen in the majority of parameters measured between the two groups excepting fluoroscopic and procedure times.

Whether or not the experienced group's performance failed to be superior to the novice's due to a true lack of difference in ability or to the simulator's inability to detect a difference is unknown. However, we presume that—due to the tremendous difference in endovascular experience and relative homogeneity in video gaming habits (Expert 50% vs. Novices 75% who played "sometimes" for Study I)—there was, in fact, a difference which was not detected. Additionally, no statistically significant difference in video gaming habits was demonstrated in Study IV in contrast to what has been found in laparoscopic VR simulator studies.[73] Given the enormous variation found in video gaming environments and individual game objectives, it would appear to be difficult to predict a significant impact on VR-Lab skills independent of the particular video games the participants regularly enjoyed.[74] That is, video game human-computer interactions vary to such an extent that some games might well contribute to VR-Lab performance (ex. flight simulators) whereas others would undoubtedly not (ex. Sims II©). Thus, the closer the visuo-spatial environment in VR resembles the real environment, the better the transferability.[47,75,76]

Alternatively, the modules in question may not have been technically challenging enough to register real skill differences. Furthermore, we noted that novices appeared very focused on giving their best performance while the experts seemed more experimental with the capabilities and limitations of the simulators. These differing attitudes, although not systematically assessed, could have contributed to the equivocal objective data. Other studies have shown very rapid learning curves for novices when their VR-Lab performances are compared to experts and the same may have been true here.[77,78] Interestingly, subjective opinions from both groups regarding the simulator's usefulness were positive, which is not an unimportant fact when assessing a new training device.[79] However, these two studies focused on the validation of the simulator's metrics rather than the transferability of subjective opinions to objective performances. Again, the focus in this dissertation was on the metrics, not the educational utility of the machine, which may very well have clinical relevance.

From the observations, it seems that basic psychomotor and procedural skills might be possible to learn using endovascular simulations. The combination of using actual, albeit modified, clinical equipment in conjunction with the Procedicus-VIST™ system's haptic feedback should, reasonably, be directly transferable to the clinical setting, i.e. the human catheterization laboratory. Procedural skill, i.e. the actual order in which the procedure in question should be performed, also seems feasible to learn outside of the catheterization laboratory. Having learned these two basic skills *ex vivo*, i.e. in an experimental environment, would allow the interventionalist to focus on the patient and the job at hand *in vivo*, i.e. in real life. While performing in the actual catheterization laboratory will always be more difficult than in the virtual one, the better one's training has been, the easier it becomes to meet real world challenges. Virtual reality training may serve as a stepping stone in attaining procedural proficiency before one proceeds to one-on-one instruction under an experienced interventionalist.

The deployment of virtual reality simulations and other forms of laboratory-based training is an adjunct, not a replacement for the traditional apprenticeship model of technical skill education. The current model suffers from a curriculum that is based on patient availability and the logging of arbitrary numbers of cases, not the achievement of specific objectives. Given the dramatic changes that are taking place world-wide with regards to work hour restrictions, systems of training that afford residents the opportunities for deliberate practice are desperately needed. The alternative might be lengthening an already too long training process, or worse, the graduation of tomorrow's specialist who are less well-trained than those of today.

There is an old adage in procedural medicine that "first you get good and then you get fast". In that vein, chronological measurements such as procedure and fluoroscopic times may be rudimentary measures of efficiency but, taken alone, are poor proxies for quality. Furthermore, exemplary results often demand *longer* procedure times and *increased* fluoroscopy to ensure exact angiographic interpretation and optimal stent placement. In contrast, a rapid yet sloppy procedure may result in iatrogenic vessel damage, cerebrovascular insult or persistent stenosis requiring open surgical repair.

General uncertainty due to such things as confusing equipment nomenclature, lack of anatomical familiarity, awkward use of the fluoroscope and complex procedure sequence may have caused the observed differences in speed between the experts and novices, rather than any appreciable differences in the final quality of the interventions. If true, a longer introduction period with the simulator and CAS training may have mitigated the observed time differences as has been found in other studies.[78,80] Interestingly, the mean differences in fluoroscopic time equates almost exactly to the greater procedure times registered in the novice group, 8.7 ($p = 0.0066$) and 8.7 ($p = 0.0031$) minutes. Therefore, it seems plausible that the only true metric performance difference lay in fluoroscopic use.

Nevertheless, fluoroscopic time measurements combined with the number of cine loops may prove to be a readily exploitable VR-lab metric. Specifically, these two metrics might serve as indirect measures for radiation exposure and lend themselves to improve a trainee's radiation discipline. That is, a trainee could spend time learning procedures in the VR-Lab and thereby decreasing real world exposure to ionizing radiation to themselves and patients while simultaneously learning better radiation protection techniques without true exposure.

Although experienced in general interventional techniques, the CAS procedure was a novel experience to the expert group. Thus, the expert group's performance may not truly represent the abilities of specialists with more CAS experience and should be considered as a plausible Type II error.

Determining the actual quality of a given interventional performance remains difficult.[81] Furthermore, standardized evaluation is complicated because critical procedural steps vary between procedure types and different simulators. Parameters such as respect for tissue, fluoroscopic proficiency, time-motion efficiency, instrument handling/safety and procedural knowledge have shown promise in assessing interventional skills transfer from the VR-Lab to the Cath-Lab.[82,83] Dayal's validation also demonstrated that GRS proctor scores were sensitive in detecting performance differences between different groups both pre- and post-training.[77] Cognitive task analysis of interventional skills currently under the guidance from the Joint Simulations Task Force (JSTF) of SIR and CIRSE promises to offer more precise protocols for use in skill assessment in the future.[84-86]

Given that carotid artery stenosis intervention is procedurally difficult, possesses an extensive learning curve and involves a grave list of potential complications, the role of non-clinical skills training is of increasing importance.[15,81] Indeed, similar arguments were given when the FDA mandated VR-Lab training in the United States prior to being certified to perform CAS stenting.[87] Thus, further validation studies are of increasing clinical importance and necessary for the development of standardized curricula.

Three parameters of clinical importance approached statistically significant levels. Experts outperformed novices with greater tool:vessel ratios and lower residual stenoses, which translates to wider dilation diameters and better post procedural end organ flow in real patients. In contrast, novices “outperformed” the experts in placement accuracy by centering their stents precisely over the carotid lesions. However, in the CAS procedure, the stent is placed over the lesion with a large portion protruding into the common carotid artery, meaning that the stent’s center will not lay over a lesion’s true center. Therefore, placement accuracy seems to be a less meaningful metric compared to the renal artery stenting where exact stent placement ensures exclusive stenting of only pathological tissue.

SKILLS TRANSFER

Many previous skills transfer studies have focused on time to completion as their major endpoint.[25,45,88] While an expert ought to be able to perform any given case faster than a novice, time to completion cannot be considered the best indicator of technical skill. The most important factor is an evaluation of the quality of the finished product.[89] Therefore, an objective method to assess procedural skills was used, i.e. Total score. This method is a deliberate, standardized complement to the mentoring method wherein an experienced master judges an apprentice to have become a safe pair of hands. Another benefit of this methodology is its stability over time. Total score skills assessment is independent of the laborious necessity of validating new simulator metrics as newer modules and metrics are developed.

What’s more, many of the experimental attempts to prove that VR-Lab acquired skills transfer either to another simulator or to the actual OR or Cath-Lab have included control cohorts which received no training.[2,12,13,39,51,52] Interestingly, the evidence from these studies has not been unanimously supportive for the transfer of VR-Lab skills to real operative theaters. Thus, at least some of the available evidence does not use an entirely fair comparison as many have tested one training form against no training at all. Study II attempted a comparison of the most advantageous training method available, i.e. using the P-Lab for an approximation of the human Cath-Lab, and produced significant results in Total scores for both forms of training and therefore supported the skills transfer hypothesis.

In the general statistical model, both training methods improved Total score significantly and with similar magnitudes for each additional session in either laboratory. This finding indicates that the two training laboratories were as effective at improving endovascular skills, yet lack of difference should not be mistaken for equality. Of note, total scores were significantly and consistently higher in the VR-Lab which confirms that the real catheterization laboratory offers a more challenging venue. An exploitation of this finding might be the substitution of performance evaluations in the virtual environment for the real catheterization laboratory. Assuming the predictive value of VR-Lab assessments, attainment of Total score benchmarks in the VR-Lab might be required prior to allowing access to perform procedures on patients.[67,78,90,91]

The data reinforces the maxim that one improves at a skill with specific practice.[4,5] Increasing number of P-Lab sessions improved P-Lab Total scores but had no influence on VR-Lab Total scores. In contrast, previous VR-Lab sessions significantly improved both P-Lab and VR-Lab Total scores. The evidence indicates that endovascular skills learned on the Procedicus-VIST™ transfer to the real catheterization laboratory-- as modeled by healthy porcine anatomy in the P-Lab. Indeed, Chaer et al, using similar performance measurements as Study II, recently demonstrated significant improvements in the real Cath-Lab in a VR trained group during proctor supervised iliac artery stenting procedures.[83]

Why P-Lab sessions failed to improve VR-Lab Total scores is of particular interest. This failure was assumed to be caused by lack of interest in the participants. In essence, the “window of opportunity” where the trainees were interested in performing well on the simulator may have passed, once they had managed the P-Lab session. If true, this may be of importance when constructing curricula as high motivation is a prerequisite for optimal training.[53,92,93] Therefore, it may be best to initially train in the VR-Lab before moving into the Cath-Lab to avoid the detrimental effects of boredom.

SUBJECTIVE EVALUATIONS OF THE VR-LAB

Subjective evaluations of the VR-Lab were favorable throughout despite observed differences between students and experts. Students’ responses were generally more positive. Of note, students expressed increased interest in endovascular specialties and further VR-Lab training at the end of the experiment suggesting that early experience in a VR-Lab may help students in making a more informed career choice.[94,95] Both groups felt the VR-Lab experience was realistic compared to their experience in the P-Lab, more effective than didactics alone and a good teaching instrument which should be mandatory in an endovascular curriculum thus recognizing that VR is a valuable adjunct to standard training.

The data indicate that most subjects would support a compulsory training curriculum combining VR-Lab and P-Lab training under the watchful eye of an experienced proctor. However, porcine experience appeared to be strongly favored over the VR simulation experience. This may be a ripple effect originating either from their low opinion of the simulator's realism or the negative influence of software "complications," i.e. computer malfunctions, experienced in the VR-Lab. We observed that trainees were keenly interested in managing OR complications and viewed them as a complement to their training. In contrast, computer "complications" had no real world value for the trainee and were viewed as an unwelcome interruption. Subjective and objective data have shown that all participants left the course with a better grasp of the IAS procedure regardless of which training group they were assigned. Lastly, in spite of the fact that most felt VR to be a good training tool, they did not think that it could replace the porcine training laboratory.

All participants felt that proctored instruction was essential and that they had a better understanding of the procedures after the study. The issue of proctoring is critical. Machines alone are not an adequate replacement for an experienced teacher. It is certainly true that most skills laboratories have found repeatedly that faculty involvement during teaching sessions is a critical element of successfully incorporating VR into procedural training. There continues to be a vital need for the continued use of the master-apprentice model in interventional skills education which has served so well in the past. While all groups grasped the advantages of VR training offered, they acknowledged the continued role of one-on-one mentoring in their path to skills perfection.

ECONOMICS

Most readers need no introduction to the impact that restrictive budgets have on the educational goals for institutions of higher learning. The demands of training increasingly complex surgical and interventional procedures have risen as the available amount of legal working hours has diminished.[54,56-58] Simply, given that two methods of training for interventional skills produce similar results, institutions should prefer the more economical method. All health care systems are experiencing financial pressures. This reality mandates that the training sector seek tools that have the optimal cost-benefit ratio. In other words, we must attempt to provide the best healthcare and medical training to the greatest number of individuals using finite resources. This philosophy is extant whether the funds available are coming from public or private sources.

Stents in the P-Lab are consumed during each procedure since a deployed stent cannot be practically retrieved. Even at volume discounted prices, this accounted for 83% of the total cost in the P-Lab. In contrast, the interventional stents used in the VR-Lab can be reused since they are "virtually" placed and constituted only 6% of the total VR-Lab costs assuming a simulator purchase, or 20% under the rental analysis. Naturally, training facilities could use older equipment from national Cath-Labs which had expired dates to offset this cost but this would lead to a course which trained novices with old, if not obsolete, techniques. Thus, the cost for an institution for a modern P-Lab stenting course would vary in proportion with the amount of stenting planned.

The purchase price of a Procedicus-VIST™ spread over a five year period as the annual depreciation value of €30,000 represented the brunt of the VR-Lab at 87% compared to the annual rental of the P-Lab at €4,140, i.e. 9% of the budget. However, this calculation assumes a residual value of the VIST™ to be €50,000 and institutional ownership of a portable VR-Lab. Perhaps of greater importance, it might be possible to upgrade the simulator with newer software packages or hardware thereby extending its life cycle beyond the five year estimate. If true, the annual depreciation value would then be modified to the new life cycle end making the true annual cost smaller than the current estimate used in the analysis. Additionally, the simulator could be rented out or used as part of a free-standing endovascular curriculum given by the purchasing institution thereby functioning as a source of potential income.

While the cost ratio from VR-Lab purchase vs. P-Lab rental analysis found the VR-Lab to be 0.75 the P-Lab course costs for the first year, the VR-Lab/P-Lab ratio in the sensitivity analysis ranged from 0.25 to 2.22, with six of the nine possible outcomes favoring the VR-Lab, Table 14. However, two of the three ratios which favored the P-Lab over the VR-Lab assumed a price increase for the VIST™. Although technical refinements and improvements of the simulator hardware and software performance may to some extent add to the cost, such a scenario can be considered highly unlikely as information technology (IT) based industries depend upon delivering more powerful computers at lower costs at regular intervals. The remaining favorable P-Lab ratio assumes a sudden decrease in the price of the P-Lab rental fee which can be considered to be unlikely. Therefore, the existing ratio or those which include a stable or falling VR-Lab estimate are the most plausible findings in the sensitivity analysis. The VR-Lab rental vs. P-Lab rental based analysis cost ratios were unanimously in favor of the VR-lab at all cost estimate fluctuations. However, as access to rental simulators was assumed to be the limiting factor for most institutions, the VR-Lab purchase vs. P-Lab rental was deemed to be the financial analysis with the most external validity, i.e. general applicability.

Finally, learning a skill and perfecting it are two different phases in the pursuit of mastery of manual skills training.[30] Repetition is an essential part of mastery, but—

presently—trainees normally use patients to perfect their skills. The natural extension of an economic analysis then becomes how many courses are required to reach a certain skill level and how do their efficacies compare, i.e. does one alternative offer quicker mastery compared to the other and, if so, how does this affect the cost ratios. Thus, not only must introductory courses be considered, but also the number of courses needed to attain benchmark levels prior to practicing in the human catheterization laboratory. In addition, refresher courses may be needed to maintain skills for low frequency procedures in clinical practice. With the current level of evidence regarding efficacy and curricula, one cannot make any meaningful economic calculations based upon required course numbers as that number is unknown.

Rapid advances in computer sciences over the last decade have heralded phenomenal technology such as virtual reality simulators capable of reproducing operative experiences outside of the OR. Although we may still be in the embryonic stages of VR development, its future appears bright. Therefore, investment in a technology that possesses as many intrinsic advantages as VR does, seems to be a wise strategy for the future. However, acceptance of the VR-lab as a venue for skills acquisition within the scientific community awaits further evidence demonstrating the transfer of skills from VR to the Cath-Lab.[32,45-47,50,83] As rational as this requirement seems, the same demands were never placed upon current non-clinical training using research or experimental animals. Remarkably, no literature supporting endovascular skills transfer from the live animal model to the OR could be found. Thus, as demonstrated by the lack of evidence regarding both labs, we still know very little about what type of training helps in the real world, Table 1.

Evidence supporting skills transfer from endoscopic and laparoscopic VR-Labs to the OR exists.[12,39,45,65,91] However, extrapolation of those results to endovascular techniques was deemed inappropriate because many of these studies compared the VR-Lab group to non-trained controls and, more importantly, lacked an endovascular focus. One study demonstrated transfer of VR-Lab acquired skills to the P-Lab.[96] No studies were found to have compared the efficacies of the P-Lab to the VR-Lab in acquiring or perfecting endovascular skills in the human catheterization laboratory.

Despite the lack of confirmative data regarding the clinical efficacy of both P-lab and VR-lab endovascular skills training, the results of the financial calculations raise an important ethical question. If skills training in the VR-lab is much less expensive than training in the P-lab, and if there is no proven superiority of using an animal lab, can skills training on animals be supported from an ethical perspective? In order to substantiate continued use of live animals, the outcome of skills training in the P-lab should be significantly more efficient than training in a VR-lab. Although further refinements of the VR-lab modules may be necessary to fully compete with P-lab training, the potential for technical development of the VR-lab suggests that the need for live animals may be abolished in the future.

ETHICS

Ziv et al produced the most recent article to bring simulations and medical ethics together in 2003.[97] The paper acknowledged that medical training, at some point, must use live patients, but that computer simulation based technology could be a valuable tool in mitigating the ethical tensions and practical dilemmas these first steps entail. Their proposed ethical framework focused on best standards of care and training, management of medical errors, respect for patient safety and autonomy and the responsible allocation of resources. Thus, the implementation of increased VR-Lab endovascular training complies with that framework and succeeds in fulfilling our ethical responsibility as researchers and educators.

Although not focused solely upon VR and ethical considerations, Gruber and Hartung pointed out that the validity of some animal research is often unknown either through lack of will or ignorance.[98] Furthermore, because the data produced using animals is commonly statistically under-powered, it is impossible to draw reliable conclusions from the observed outcomes. The parallel between Gruber and Hartung's observations and the uncertainty surrounding P-Lab skills transfer efficacy are not difficult to recognize. Even though the P-Lab mimics the OR and involves the use of real interventional equipment to develop skills, it does so on non-human species with healthy anatomy and produces unknown results in the trainee. Studies need to be performed to quantitatively measure the improvements that have, thus far, been assumed by training on pigs. The same arguments hold true for the VR-Lab. As the experts in this field have recently pointed out, "Intuition is not evidence."[50] Yet, neither does lack of evidence mean that something is untrue, merely unproven.

On the other hand, the possibility of training a technique incorrectly and/or building false confidence exists when using an untried method. What if one or both of the labs teach, unknowingly, a skill that will be *harmful* to cath-lab performance? One cannot assume that these training forms offer only benefit; it must be proven. Also, while mechanical skills can be taught given the proper setting, most experienced interventional radiologist will agree that clinical judgment and situational awareness are the hardest to teach. Specifically, overstepping one's ability due to overconfidence and ignoring common problems due to ignorance seems to plague beginners.

Despite the fact that the general public's emphasis on ethical responsibility commonly outweighs economical issues, pragmatic use of finances is necessary. Gruber and Fitzpatrick state that ethical considerations alone will do little to bring about change.[98] Furthermore, they remind us of our responsibility to ask only important questions and to design experiments with ethics in mind which minimize suffering.[98,99]

According to Machan, animals lack rights because they do not have the faculty to make moral choices.[100] Thus, their unaccountable actions are a consequence of instincts rather than conscious deliberation between what is right or wrong. He reminds us that humans are also a part of nature and, as such, make use of the animals below us to survive and thrive, as do all of nature's creatures. However, even though the lesser animals lack such moral agency, we are bound by ethics to treat them with consideration. In essence, we ought to be thinking of a better way to get the experimental data humanity needs while including the three R's, i.e. 1) Replacement 2) Reduction and 3) Refinement, at each step.

Thus, the main argument for P-Lab training is to increase patient safety during the steepest part of the learning curve. Improvement in endovascular skills does not appear to be part of the argument. In fact, the efficacy of P-Lab training on human catheterization laboratory is unproven placing it on a level of uncertainty equal to the VR-Lab. Therefore, two alternative methods of training endovascular novices are available, which offer the benefit of early learning curve safety to patients. The choice of which alternative to use remains unanswered yet the default remains the P-Lab in countries where it is available.

Thus far, we have dealt with skills training removed from the real catheterization laboratory in order to increase patient safety and save costs. Yet our most common training venue remains the human cath-lab. While skills transfer from the human cath-lab is a moot point, the ethics of using patients for skill acquisition and practice remains problematic. Is it ethical to include humans in our training? The benefit to the patient—and potential harm—must be weighed against the student' benefit as they acquire experience. Even though the use of research animals raises ethical concerns as discussed previously, should "skills training" be considered research or merely a form of consumption? Research animals are used in the P-Lab to gain practical experience removed from the human laboratory, not to increase scientific knowledge. We train non-clinically that we might spare fellow humans harm caused by our lack of experience. Yet, can we make the decision to expend another sentient being without really knowing if it serves its stated purpose? The addition of validated training method(s) would certainly be welcomed by fully informed, consenting patients prior to becoming a fellow's first stenting procedure. Furthermore, an adequately trained fellow might enter the human laboratory at a higher point on their learning curve with a decreased likelihood of errors and faster procedure times, both of which lead to cost reduction.[65]

STUDY LIMITATIONS

First, the sizes of the studies are a limiting factor. In the validation studies I and IV, we were unable to recruit sufficient numbers of “unexposed” novices and interventionalists to substantiate the lack of evidence. However, post-hoc analysis for the two key metrics of residual stenosis and placement accuracy revealed that the numbers needed were not possible to reach within Sweden. What’s more, the use of the simulator as a proficiency assessment instrument is, at best, an “off-label” use as it was designed to educate not regulate. The final approval of a novice’s clinical proficiency ought to lie in the hand of an endovascular expert with years of clinical experience, and not in the hands a software programmer.

Study II was also small due to the logistical and financial limitations of the laboratories used. Nevertheless, the study design did prove valuable in detecting a transferability of virtually generated skills. The measured skill improvements may not represent the maximized training benefit of either method. A fairer comparison would have included longer training periods in both laboratories. The small but significant gains observed may lack clinical relevance. In the general statistical model, either training form improved Total score by only 6.8% of the total maximum possible per session. Lastly, the significance level for the critical result, i.e. VR-Lab sessions improved P-Lab performance, lay just within the acceptable $p = 0.05$ level. Stronger significance would have enhanced the trustworthiness of our findings.

Any financial analysis is rift with assumptions, which are presumed to be correct but are often only educated guesses based on local and historical financial data. Thus, despite including sensitivity analysis for both the VR-Lab purchase/rental vs. P-Lab rental, real economic data could vary greatly not only from country to country but also between institutions within the same metropolis. Therefore, the true external validity of the findings, although likely to be accurate in some situations, could be unrealistic in others.

Finally, these projects have included only one type of endovascular simulator. Several are available on the open market, but our laboratory only had access to one due to funding issues. The conclusions drawn from the data relate only to this particular simulator and not to VR simulations in general. Better methodology would have included each of the existing simulators in their own experimental group to detect possible differences between the models as has been done for other specialties.[101]

FUTURE RESEARCH

The true litmus test of this simulator's usefulness will lie in its ability or failure to significantly improve human catheterization laboratory performances vis-à-vis improved objective performance scores significantly related to VR training. However, definitive proof for the reduction of intraoperative errors would also bolster support for the incorporation of skills training in the VR-Lab. To that end, a human transfer study is in progress at the Interventional Radiology Department, Kuopio University, Finland using a similar clinical assessment scale as the Total Score used in Study II, renamed the Objective Assessment of Scandinavian Interventional Skills (OASIS). Results are expected by the end of 2007.

Furthermore, the rapid advances in computer technology all but guarantees periodic software updates to the metrics of this and other simulator systems. Naturally, once a new set of metrics are introduced, they too will need to be systematically validated. As most validation studies consist of relatively small n-values—our studies included—confidently claiming or denouncing construct validity will continue to be based upon lower powered studies. In order to automate the collection and analysis of larger data sets, a kiosk database program has been developed which could be delivered to high volume VR training centers worldwide such as Abbott Vascular's Crossroads training institute in Diegem, Belgium. Software-version-specific metric data can, in this fashion, be collected in conjunction with demographic/IR experience data from the different locations. Periodic downloading to a central validation authority, such as the Joint Simulation Task Force (JSTF) would greatly speed the validation process and increase our confidence in the use of VR technology.

Yet another application of simulation technology is the ability to perform "mission rehearsals", termed "procedure rehearsal" within medicine, for critical procedures such as neuroradiological interventions. In collaboration with Anthony Gallagher, we have initiated a live case procedure rehearsal study scheduled for May 2007. The participants include neuroradiologists from Turkey and four patients, two of which will have had their CT scans converted into patient specific VR modules on the Procedicus-VIST™. The potential benefits of being able to perform a clinical case in the Cath-Lab which one has practiced to perfection in the virtual environment may prove to be of increasing clinical importance in the future.

Lastly, the previously mentioned JSTF has launched a detailed task analysis of many IR procedures. The major aim of this collaboration is to dissect complex procedures down to their basic steps with the help of cognitive behavioral psychologists. These steps will then be weighted according to their importance to the end quality of the intervention in question. In this way, a higher quality skills assessment scale is being devised to produce clinically relevant performance measurements and potential new metric parameters. The results of this effort will finally give the endovascular community their gold standard. More importantly, the information could be given openly to the industry to stimulate the development of clinically significant metrics for future VR simulators. In this way, the medical community could step forward to take a proactive role in developing future medical simulators versus waiting for the industry to independently create them in the absence of guidance as to what is important to the end users.

Although currently useful for endovascular training, further improvements in the VR-Lab are needed for this training form to reach its full potential as a true full-scale interventional simulator. Namely, the inclusion of interventional and clinical complications such as sub-intimal catheterizations, iatrogenic thromboembolism to end organs, coronary arrythmias/infarction, apoplexia, acute reactions to contrast medium and vessel puncture complications.. Furthermore, femoral artery puncture, i.e. the very start of a procedure, constitutes a substantial obstacle for beginners. There is always a risk for dissection, embolus, thrombosis, development of hematoma and pseudoaneurysms after an arterial puncture. Likewise, arteriotomy closure, using various closure device instruments presents another application which lends itself to virtual reality training in order to avoid or minimize the incidence of post-operative complications such as pseudo aneurysms, inguinal hematomas and fistula formations. In order for such developments to take place, the medical community should initiate or increase their contact and cooperation with the industry.

CONCLUSIONS

I. With the exception of total procedure times, the renal artery stenosis modules failed to demonstrate construct validity. A) Despite demonstrating face validity and some non-significant absolute performance metric differences between groups, the virtual reality simulator was unable to stratify virtual interventional performances based upon experience. B) Current renal module metric parameters were not capable of performance assessment outside the catheterization laboratory. C) The VR-Lab was generally approved as a pedagogic tool based upon the subjective data.

II. Porcine and virtual reality laboratory training appears to produce comparable training results in the endovascular novice as measured in the porcine laboratory. A) Virtual reality simulation training was found to be as effective as the porcine laboratory in improving P-Lab iliac artery stenting performance. B) Skills learned in virtual reality using the iliac artery modules may transfer to the catheterization laboratory as simulated by healthy porcine anatomy. C) While both laboratories were subjectively approved for endovascular skills education, the P-Lab was apparently favored by trainees based upon subjective data.

III. An economical analysis of using virtual reality simulations versus the porcine laboratory to train endovascular skills was performed. A) The purchase of a simulator cost less compared to renting the P-Lab in the short and long term perspective. B) Rental of a VR-Lab was found to be considerably less expensive than renting the P-Lab to meet national training demands, although locations without adequate rental access would make such an implementation difficult.

IV. With the exception of procedure and fluoroscopic times, the carotid artery modules failed to demonstrate construct validity. A) Despite demonstrating face validity, the simulator metrics were unable to stratify virtual interventional performances based upon prior endovascular experience. B) The Procedicus-VIST™ metrics did not work as an assess tools for endovascular skills outside of the catheterization laboratory. C) The VR-Lab was subjectively approved as a pedagogic tool.

SVENSK SAMMANFATTNING

BAKGRUND: Interventionell radiologi av idag innebär diagnostik och behandling av sjukdomar med hjälp av perkutana minimalinvasiva, oftast kateterstyrda procedurer med stöd av bildgivande radiologiska metoder (datortomografi, magnetkamera, röntgengenomlysning, angiografi, ultraljud). Endovaskulär radiologi är en förhållandevis ung gren av medicinen och etablerades först av pionjärer på 1960-talet. Den första perkutana transluminala kateterbehandlingen rapporterades av Dotter i femoro-politeala kärlområdet i *Circulation* 1964. Tekniken vidareutvecklades av Gruntzig till att innefatta en angioplastikballong och via Palmaz introducerades på 80-talet stentbehandlingar av blodkärl. Liksom inom kirurgin har interventionell radiologi anammat samma utbildningskoncept, dvs. att överföringen av kunskap ska ske i ett mästar-lärlingsförhållande där röntgenrummet med angiografiutrustning utgör klassrummet. Basal endovaskulär teknik har därför utförts på människa direkt, även om detta inte varit direkt utsagt.

Det finns många faktorer som talar för att den basala träningen av interventionell radiologi ska flyttas ut från den direkta patientkontakten i operationsrummet: För det första har en rapport från USA (Kohn, Corrigan, Donaldsson, 1999) påvisat att mellan 44000 och 98000 individer per år avlider till följd av misstag inom sjukvården. Fel eller misstag inom sjukvården kan indelas i systemfel och i individuella fel. Analys och tillfälle till reflektion möjliggör att vi kan lära av våra misstag. Det måste då finnas säkerhetssystem som objektivt kan utvärderas. Kvalitetssäkring av utbildning, vidareutbildning och kompetensbevarande utbildning är nödvändigt. Ahlberg (Thesis Ahlberg, 2005) har kritiserat den äldre mästar-lärlingfunktionen bl.a. därför att den är svår att utvärdera.

För det andra är bildstyrd intervention där operatören vägleds av en tvådimensionell monitorbild av en patients inre komplex. Metoden kan anses vara relativt sett svårare att lära än öppen kirurgi. Höga krav ställs på visuell och spatial förmåga. Den ökade komplexiteten innebär också en ökad risk för komplikationer.

För det tredje har interventionell radiologi de senaste åren fått en mycket stor och ökande betydelse som alternativ och komplement till sedvanlig kirurgisk behandling av sjukdomar i kärlsystemet. Genom dess minimalt invasiva karaktär har dessa interventionella metoder lett till att patientgrupper som tidigare inte kunnat behandlas på grund av risk för komplikationer, nu kan åtgärdas med framgång och med betydligt mindre risker.

Komplicerade nya endovaskulära tekniker har införts, t.ex. endovaskulär aortareparation, EVAR. De kärlproteser som finns kräver utbildning i handhavande och certifiering och alltmer komplicerad teknologi är på väg. Risken för komplikationer ökar med ökande komplexitet.

För det fjärde finns det i all färdighetsutövande en inlärningskurva innan man uppnått en godtagbar standard. De flesta svårigheter och komplikationer som inträffar, sker tidigt i inlärningsfasen men bidrar också till kunskapsinhämtande. Att flytta träning från angiosalen till en simulator skulle kunna förbättra säkerheten, då de initiala momenten kan övas till dess de integreras som en färdighet (jämför att lära sig cykla, köra bil). Träning av medicinska färdigheter samt träning av kommunikation, ledarskap och samarbete i simulatormiljö bör leda till en ökad patientäkerhet. Misstag kan minimeras om medicinska färdigheter och kompetenser kan öka i simulatormiljö.

För det femte är en grundläggande färdighet i punktionsteknik och invasiv angiografi en del i att utvecklas till en duktig interventionalist. Antalet utförda angiografier som krävs innan en operatör kan utföra en angioplastik under handledning är en bedömningsfråga. De enkla diagnostiska angiografierna har nu minskat dramatiskt och möjligheterna för unga läkare att få en basal färdighet har därmed minskat. Alternativa metoder är träning på djur/kadaver för att skona patienter från otränad personal eller att öva med medicinska simulatorer.

Syftet med denna avhandling har varit:

att undersöka betydelsen av simuleringsteknologi med avseende på validering av en simulator för kateterburen diagnostik och intervention

att jämföra endovaskulär träning på simulatorn med träning på djur

att genomföra en ekonomisk analys av kostnaderna för simulatorträning med kostnaderna för träning på gris.

I DELARBETE I var målsättningen att studera om simulatorn kunde diskriminera mellan erfarna radiologer (n=8) och läkarstudenter (n=8) på termin VIII, med hjälp av de metriska resultat som simulatorn levererar. Alla deltagarna fick en kort introduktion i endovaskulär teknik och i handhavandet av simulatorn. Därefter fick alla deltagarna på förmiddagen arbeta med sex olika patientfall med njurartärförträngning. På eftermiddagen registrerades deltagarnas resultat i en ny omgång på alla sex fallen. Det visade sig att studenterna använde längre genomlysningstid (mer röntgenstrålning) än erfarna läkare, men i övrigt skiljde sig inte resultaten signifikant åt. Någon värdering av deltagarnas färdigheter av en erfaren radiolog gjordes inte i denna studie. Både studenter och erfarna radiologer bedömde att simulatorn är ett viktigt pedagogiskt instrument och att den är en god spegling av verkligheten (face validity).

I DELARBETE II deltog tolv läkare med varierande erfarenhet i interventionell radiologi i en tvådagars kurs för att utveckla sin teknik i endovaskulär stentning av bäckenartärer. Deltagarna indelades i 4 grupper. Med hjälp av en utsänd enkät lottades deltagarna via ett statistiskt förfarande så att kunskapsnivån i de 4 grupperna skulle kunna vara så lika som möjligt med avseende på erfarenhet. Den första gruppen fick enbart träna endovaskulär teknik på gris, den fjärde gruppen enbart på en medicinsk simulator under de två kursdagarna. Mellangrupperna fick träna på gris första dagen och simulator dag två. Deltagarnas basala kunskapsnivåer värderades av särskilt utsedda erfarna lärare både på gris och simulator vid fyra tillfällen, d.v.s. på morgonen och vid dagens slut båda kursdagarna. Ett särskilt utvärderingsprotokoll modifierat från Dr Reznick, Professor, Toronto, Canada och Dr Beard, Consultant Vascular Surgeon, Leicester, England användes. Videoinspelning av deltagarnas prestationer utvärderades separat av två mycket erfarna specialister i ett senare skede. Simulatorträning och gristräning visade sig båda vara effektiva i att förbättra den endovaskulära tekniken. Den var en god samstämmighet mellan videobedömning och bedömning på plats av deltagarnas insatser. Träning i simulatormiljö medförde att deltagarna presterade bättre vid den därpå följande träningen på gris (som ska efterlikna patientsituationen). Båda träningsmetoderna ansågs av kursdeltagarna vara goda miljöer för träning.

I DELARBETE III analyserades de ekonomiska kostnaderna i studie II. Två scenarier analyserades: Först jämfördes kostnaden att köpa en simulator med att hyra ett gristräningslaboratorium under fem år. Denna analys ansågs mest realistisk, då många institutioner idag har ett djurlaboratorium, men få institutioner äger en simulator. Sist jämfördes också kostnader för att hyra simulatorn med att hyra ett djurlaboratorium under samma period, d.v.s. fem år. De faktiska kostnaderna för vårt försök i delarbete 2, liksom de båda beskrivna scenarierna visade att en medicinsk simulator är betydligt mer ekonomiskt fördelaktig jämfört med träning på gris. Sannolikt blir simulatorerna allt mer sofistikerade och billigare i framtiden. Etiskt framstår medicinska simulatorer som ett självklart förstahandsval för träning av endovaskulära procedurer.

I DELARBETE IV deltog 16 läkare med varierande erfarenhet av interventionell radiologi och 16 läkarstudenter från den avslutande delen av sin utbildning. En komplicerad endovaskulär procedur, stentning med angioplastik av arteria carotis interna (halspulsådern) med filterskydd, introducerades för deltagarna. Ingen av deltagarna hade utfört en sådan procedur på människa. I denna studiedesign fick kursdeltagarna en kort introduktion av proceduren och i handhavandet av simulatorn. Deltagarna fick sedan öva på ett fall, därefter värderade simulatorn resultaten på nästa fall. Det visade sig att den totala behandlingstiden och röntgengenomlysningstiden var signifikant kortare för de erfarna läkarna jämfört med studenterna. De båda grupperna angav att simulatorn är ett bra pedagogiskt verktyg och också efterliknar verkligheten väl.

SAMMANFATTNING: Våra resultat visar att en endovaskulär simulator är ett betydelsefullt pedagogiskt instrument och också en god spegelbild av den endovaskulära miljön, samt ett ekonomiskt fördelaktigt alternativ till träning på djur. Simulatorer kan dock inte idag, med de återkopplingsinstrument (metrics) som finns tillgängliga, visa skillnaden mellan erfarna och oerfarna operatörer, mer än i tidshänseende och i användandet av röntgengenomlysning.

Framtiden: Man bör i framtida studier analysera vilka moment i en endovaskulär procedur som är viktiga att lära sig och hur komplikationer och bra respektive dålig teknik skall defineras. Denna kunskap bör sedan implementeras i simulatorernas återkopplingssystem (metrics). Kanske kan då eleverna träna ensamma på simulatorerna tills de uppnår en viss färdighet. För att få ut så mycket som möjligt av träning med en endovaskulär simulator bör man också integrera flera utbildningsaspekter i ett curriculum. Träningen bör ses ur ett helhetsperspektiv. Att optimera utbildningen avseende självstudier, föreläsningar om proceduren, träning i simulator med och utan handledning och på sikt kunna definiera en certifieringsnivå för utförande av procedurer på patienter är utmaningar för framtida forskning.

ACKNOWLEDGEMENTS

Richard Reznick, MD, M.Ed., FRCSC, FACS, Professor and Chair, Department of Surgery, University of Toronto for his enormous contributions and support from abroad.

Ted Lystig, Ph.D., Biostatistician, AstraZeneca; a selfless and congenial biostatistician of unrivaled caliber.

Jonathan Beard, M.B.B.S., M.Ed., FRCS, Department of Vascular Surgery, Sheffield, United Kingdom whose generosity and encouragement has been priceless.

Hans Klingenstierna, M.D., Department of Interventional Radiology, Södra Älvsborg Sjukhus, Boras, Sweden for his directness, brilliant work ethic and almost permanent smile.

Jan Göthlin, M.D., Ph.D., Professor Emeritus, Department of Radiology, The Sahlgrenska Academy at Göteborg University, Sweden for inspiring me to dig deeper and fight the good fight.

Mikael Hellström, MD, PhD, Professor and Chair, Department of Radiology, The Sahlgrenska Academy at Göteborg University, Sweden for allowing this dissertation to be undertaken within his department and his tireless support through the highs and lows.

Finally, I want to express my gratitude for my academic supervisor/mentor/trouble-maker, Lars Lönn. Without your contagious energy and inspiration, I'd be much better rested (yet unenlightened). Deepest thanks.

Academic Supervisor, Lars Lönn, M.D., Ph.D., Associate Professor, Department of Radiology, Sahlgrenska University Hospital, The Sahlgrenska Academy at Göteborg University, Sweden and Department of Radiology, Rigshospitalet, Copenhagen, Denmark

Department of Radiology, Sahlgrenska University Hospital, in particular the Uro-Gastro Section without whom, Paper II would have never been possible.

Department of Interventional Radiology, Rigshospitalet, Copenhagen Denmark, for the opportunity to gain some clinical experience under their expert tutelage.

Robert Emilsson, Photographer, The Sahlgrenska Academy at Göteborg University, Sweden for expert videographic contributions.

Professor Mats Sandberg, Ph.D., Department of Biochemistry, The Sahlgrenska Academy at Göteborg University, Sweden for introducing me to scientific research.

Professor Keiko Funa, M.D., Ph.D., Department of Biochemistry and Cellbiology, The Sahlgrenska Academy at Göteborg University, Sweden for allowing me to test the waters of research in her prestigious laboratory.

Erney Mattson, M.D., Ph.D., The Wallenberg Laboratory, The Sahlgrenska Academy at Göteborg University, Sweden for his insightful and beneficial suggestions.

Dan Lukes, M.D., Ph.D., The Wallenberg Laboratory, The Sahlgrenska Academy at Göteborg University, Sweden for his kind words of encouragement and guidance.

Lena Björneld, Department of Radiology, The Sahlgrenska Academy at Göteborg University, Sweden for her expert editorial assistance.

Ingela Björholt, PhD., Nordic Health Economic Research AB, Sweden for help in sorting out the details of the financial analysis.

Lotta Robertsson, The Sahlgrenska Academy at Göteborg University, Sweden for always having patience and a few kind words to spare.

Mentice Medical Simulations, Gothenburg, Sweden for magnitudes of cooperation and their desire to make simulator technology mainstream.

Experimental Biomedical Institute especially Eva Oscarsson and her team.

Bard Nordic AB, Martin Bengtsson and Lotta Schrewelius.

Cordis AB, Johnson and Johnson, Daniel Karlsson and Henrik Nyman.

General Electric AB, Sweden

Philips AB, Sweden

Cardiovascular and Interventional Radiological Society of Europe, CIRSE, Educational Research Grant 2005

Göteborgs Läkarsällskap for graciously funding portions of this work.

Svensk Förening för Medicinsk Radiologi

I would also like to express my gratitude to the following for their unconditional and untiring support:

Sidsel Berry
Maw and Paw
Inge and Berit Danielsson
Jimmy and Helle Witved
Martin Carlwe, M.D.
Shahin Mohseni, M.D.
Malin Lönn, Ph.D.

Thanks for putting up with me,

Max

REFERENCES

---

1. **Lobato AC, Rodriguez-Lopez J, Diethrich EB. Learning curve for endovascular abdominal aortic aneurysm repair: evaluation of a 277-patient single-center experience. *J Endovasc Ther* 2002;9(3):262-8.**
2. **Ahlberg G, Hultcrantz R, Jaramillo E, Lindblom A, Arvidsson D. Virtual reality colonoscopy simulation: a compulsory practice for the future colonoscopist? *Endoscopy* 2005;37(12):1198-204.**
3. **Piechaud PT, Pansadoro A. Transfer of skills from the experimental model to the patients. *Curr Urol Rep* 2006;7(2):96-9.**
4. **Tassios PS, Ladas SD, Grammenos I, Demertzis K, Raptis SA. Acquisition of competence in colonoscopy: the learning curve of trainees. *Endoscopy* 1999;31(9):702-6.**
5. **Parry BR, Williams SM. Competency and the colonoscopist: a learning curve. *Aust N Z J Surg* 1991;61(6):419-22.**
6. **Champion HR, Gallagher AG. Surgical simulation - a 'good idea whose time has come'. *Br J Surg* 2003;90(7):767-8.**
7. **Vozenilek J, Huff JS, Reznek M, Gordon JA. See one, do one, teach one: advanced technology in medical education. *Acad Emerg Med* 2004;11(11):1149-54.**
8. **Ziv A, Ben-David S, Ziv M. Simulation based medical education: an opportunity to learn from errors. *Med Teach* 2005;27(3):193-9.**
9. **Gallagher AG, Cates CU. Virtual reality training for the operating room and cardiac catheterisation laboratory. *Lancet* 2004;364(9444):1538-40.**
10. **Madan AK, Frantzides CT, Sasso LM. Laparoscopic baseline ability assessment by virtual reality. *J Laparoendosc Adv Surg Tech A* 2005;15(1):13-7.**
11. **Madan AK, Frantzides CT, Tebbit C, Quiros RM. Participants' opinions of laparoscopic training devices after a basic laparoscopic training course. *Am J Surg* 2005;189(6):758-61.**
12. **Youngblood PL, Srivastava S, Curet M, Heinrichs WL, Dev P, Wren SM. Comparison of training on two laparoscopic simulators and assessment of skills transfer to surgical performance. *J Am Coll Surg* 2005;200(4):546-51.**

13. Hyltander A, Liljegren E, Rhodin PH, Lonroth H. The transfer of basic skills learned in a laparoscopic simulator to the operating room. *Surg Endosc* 2002;16(9):1324-8.

14. Korndorffer JR, Jr., Dunne JB, Sierra R, et al. Simulator training for laparoscopic suturing using performance goals translates to the operating room. Transfer of skills from the experimental model to the patients Training in laparoscopic urology. *J Am Coll Surg* 2005;201(1):23-9.

15. Lin PH, Bush RL, Peden EK, et al. Carotid artery stenting with neuroprotection: assessing the learning curve and treatment outcome. *Am J Surg* 2005;190(6):850-7.

16. 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(10):1514-8.

17. Hammond I, Karthigasu K. Training, assessment and competency in gynaecologic surgery. *Best Pract Res Clin Obstet Gynaecol* 2005.

18. Schijven MP, Berlage JT, Jakimowicz JJ. Minimal-access surgery training in the Netherlands: a survey among residents-in-training for general surgery. *Surg Endosc* 2004;18(12):1805-14.

19. Cosman PH, Cregan PC, Martin CJ, Cartmill JA. Virtual reality simulators: Current status in acquisition and assessment of surgical skills. *ANZ Journal of Surgery* 2002;72(1):30-34.

20. Madan AK, Frantzides CT, Shervin N, Tebbit CL. Assessment of individual hand performance in box trainers compared to virtual reality trainers. *Am Surg* 2003;69(12):1112-4.

21. Mantovani F, Castelnuovo G, Gaggioli A, Riva G. Virtual reality training for health-care professionals. *Cyberpsychol Behav* 2003;6(4):389-95.

22. Gallagher AG, McClure N, McGuigan J, Crothers I, Browning J. Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer virtual reality (MIST VR). *Endoscopy* 1999;31(4):310-3.

23. Colt HG, Crawford SW, Galbraith O, 3rd. Virtual reality bronchoscopy simulation: a revolution in procedural training. *Chest* 2001;120(4):1333-9.

24. Gor M, McCloy R, Stone R, Smith A. Virtual reality laparoscopic simulator for assessment in gynaecology. *Bjog* 2003;110(2):181-7.

25. **Carter FJ, Schijven MP, Aggarwal R, et al. Consensus guidelines for validation of virtual reality surgical simulators. *Surg Endosc* 2005;19(12):1523-32.**

26. **Piaget J. The stages of the intellectual development of the child. *Bull Menninger Clin* 1962;26:120-8.**

27. **Jandering A. Virtual Skill Training: A Learning Theory Approach. Masters Thesis: Chalmers University of Technology, 2006.**

28. **Vroom VH. Work and motivation. New York: John Wiley, 1964.**

29. **Barnes RW. Surgical handicraft: teaching and learning surgical skills. *Am J Surg* 1987;153(5):422-7.**

30. **Reznick RK, MacRae H. Teaching surgical skills--changes in the wind. *N Engl J Med* 2006;355(25):2664-9.**

31. **Dankelman J, Wentink M, Grimbergen CA, Stassen HG, Reekers J. Does virtual reality training make sense in interventional radiology? Training skill-, rule- and knowledge-based behavior. *Cardiovasc Intervent Radiol* 2004;27(5):417-21.**

32. **Gallagher AG, Ritter EM, Champion H, et al. Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 2005;241(2):364-72.**

33. **Sherman V, Feldman LS, Stanbridge D, Kazmi R, Fried GM. Assessing the learning curve for the acquisition of laparoscopic skills on a virtual reality simulator. *Surg Endosc* 2005;19(5):678-82.**

34. **Brunner WC, Korndorffer JR, Jr., Sierra R, et al. Laparoscopic virtual reality training: are 30 repetitions enough? *J Surg Res* 2004;122(2):150-6.**

35. **Gallagher AG, Satava RM. Virtual reality as a metric for the assessment of laparoscopic psychomotor skills. Learning curves and reliability measures. *Surg Endosc* 2002;16(12):1746-52.**

36. **Gallagher AG, Lederman AB, McGlade K, Satava RM, Smith CD. Discriminative validity of the Minimally Invasive Surgical Trainer in Virtual Reality (MIST-VR) using criteria levels based on expert performance. *Surg Endosc* 2004;18(4):660-5.**

37. **Gallagher AG, Richie K, McClure N, McGuigan J. Objective psychomotor skills assessment of experienced, junior, and novice laparoscopists with virtual reality. *World J Surg* 2001;25(11):1478-83.**

38. **Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997;84(2):273-8.**

39. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg* 2004;91(2):146-50.

40. Faulkner H, Regehr G, Martin J, Reznick R. Validation of an objective structured assessment of technical skill for surgical residents. *Acad Med* 1996;71(12):1363-5.

41. Beard JD, Jolly BC, Newble DI, Thomas WE, Donnelly J, Southgate LJ. Assessing the technical skills of surgical trainees. *Br J Surg* 2005;92(6):778-82.

42. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. Testing technical skill via an innovative "bench station" examination. *Am J Surg* 1997;173(3):226-30.

43. MacRae H, Regehr G, Leadbetter W, Reznick RK. A comprehensive examination for senior surgical residents. *Am J Surg* 2000;179(3):190-3.

44. Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med* 1998;73(9):993-7.

45. Lehmann KS, Ritz JP, Maass H, et al. A prospective randomized study to test the transfer of basic psychomotor skills from virtual reality to physical reality in a comparable training setting. *Ann Surg* 2005;241(3):442-9.

46. Rotnes JS, Kaasa J, Westgaard G, et al. A tutorial platform suitable for surgical simulator training (SimMentor). *Stud Health Technol Inform* 2002;85:419-25.

47. Strom P, Kjellin A, Hedman L, Wredmark T, Fellander-Tsai L. Training in tasks with different visual-spatial components does not improve virtual arthroscopy performance. *Surg Endosc* 2004;18(1):115-20.

48. Di Lorenzo N DJ. Surgical Training and Simulations. *Business briefing: Global Surgery--Future Directions* 2005.

49. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surg Endosc* 2003;17(10):1525-9.

50. Gould DA, Reekers JA, Kessel DO, et al. Simulation devices in interventional radiology: caveat emptor. *Cardiovasc Intervent Radiol* 2006;29(1):4-6.

51. Torkington J, Smith SG, Rees BI, Darzi A. Skill transfer from virtual reality to a real laparoscopic task. *Surg Endosc* 2001;15(10):1076-9.

52. Ahlberg G, Heikkinen T, Iselius L, Leijonmarck CE, Rutqvist J, Arvidsson D. Does training in a virtual reality simulator improve surgical performance? *Surg Endosc* 2002;16(1):126-9.

53. Schijven MP, Jakimowicz JJ, Broeders IA, Tseng LN. The Eindhoven laparoscopic cholecystectomy training course--improving operating room performance using virtual reality training: results from the first E.A.E.S. accredited virtual reality trainings curriculum. *Surg Endosc* 2005;19(9):1220-6.

54. Representatives USSaUSHo. Patient and Physician Safety and Protection Act of 2005. S.1297 / H.R.1228, 2005.

55. ACGME ACoGME. Common Program Requirements, 2003.

56. AAMC AoAMC, ed. AAMC Policy Guidance on Graduate Medical Education: Assuring Quality Patient Care and Quality Education, 2001.

57. AMA AMA. H-310.979 Resident Physician Working Hours and Supervision. 2005.

58. EU TCotEU. Working Time Directive, 1993.

59. Russell WMS, Burch RL. The principles of humane experimental technique. London: Methuen, 1959.

60. Orlans FB. Ethical decision making about animal experiments. *Ethics Behav* 1997;7(2):163-71.

61. Bernstein MH. Without a tear : our tragic relationship with animals. Urbana: University of Illinois Press, 2004.

62. DeGrazia D. The ethics of animal research: what are the prospects for agreement? *Camb Q Healthc Ethics* 1999;8(1):23-34.

63. Regan T. The case for animal rights. Updated with a new preface ed. Berkeley: University of California Press, 2004.

64. Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. *Am J Surg* 2003;185(2):146-9.

65. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg* 2002;236(4):458-63; discussion 463-4.

66. Strom P, Kjellin A, Hedman L, Johnson E, Wredmark T, Fellander-Tsai L. Validation and learning in the Procedicus KSA virtual reality surgical simulator. *Surg Endosc* 2003;17(2):227-31.

67. Gallagher AG, Smith CD, Bowers SP, et al. Psychomotor skills assessment in practicing surgeons experienced in performing advanced laparoscopic procedures. *J Am Coll Surg* 2003;197(3):479-88.

68. NIH. Cost-Benefit Analysis Guide for NIH IT Projects. In: Services HaH, ed: Deputy Chief Information Officer, Center for Information Technology, 1999.

69. R: A Language and Environment for Statistical Computing [program]. Vienna, Austria, 2005.

70. Zeger SL, Liang, K.Y. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986(42):121-30.

71. Grantcharov TP, Rosenberg J, Pahle E, Funch-Jensen P. Virtual reality computer simulation. *Surg Endosc* 2001;15(3):242-4.

72. National Institutes of Health (U.S.). NIH Guide for Grants and Contracts, Vol. 14, No. 8, 25 Jun 1985. Special Edition : Laboratory Animal Welfare. Bethesda, Md.: National Institutes of Health, 1985.

73. Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Impact of hand dominance, gender, and experience with computer games on performance in virtual reality laparoscopy. *Surg Endosc* 2003;17(7):1082-5.

74. McClusky DA, 3rd, Ritter EM, Lederman AB, Gallagher AG, Smith CD. Correlation between perceptual, visuo-spatial, and psychomotor aptitude to duration of training required to reach performance goals on the MIST-VR surgical simulator. *Am Surg* 2005;71(1):13-20; discussion 20-1.

75. Arora H, Uribe J, Ralph W, et al. Assessment of construct validity of the endoscopic sinus surgery simulator. *Arch Otolaryngol Head Neck Surg* 2005;131(3):217-21.

76. Enochsson L, Westman B, Ritter EM, et al. Objective assessment of visuospatial and psychomotor ability and flow of residents and senior endoscopists in simulated gastroscopy. *Surg Endosc* 2006.

77. Dayal R, Faries PL, Lin SC, et al. Computer simulation as a component of catheter-based training. *J Vasc Surg* 2004;40(6):1112-7.

78. Hsu JH, Younan D, Pandalai S, et al. Use of computer simulation for determining endovascular skill levels in a carotid stenting model. *J Vasc Surg* 2004;40(6):1118-25.

79. Blyth P, Anderson IA, Stott NS. Virtual reality simulators in orthopedic surgery: what do the surgeons think? *J Surg Res* 2006;131(1):133-9.

80. Aggarwal R, Black SA, Hance JR, Darzi A, Cheshire NJ. Virtual reality simulation training can improve inexperienced surgeons' endovascular skills. *Eur J Vasc Endovasc Surg* 2006;31(6):588-93.

81. Gallagher AG, Cates CU. Approval of virtual reality training for carotid stenting: what this means for procedural-based medicine. *Jama* 2004;292(24):3024-6.

82. Berry M, Lystig T, Reznick R, Lonn L. Assessment of a virtual interventional simulator trainer. *J Endovasc Ther* 2006;13(2):237-43.

83. Chaer RA, Derubertis BG, Lin SC, et al. Simulation improves resident performance in catheter-based intervention: results of a randomized, controlled study. *Ann Surg* 2006;244(3):343-52.

84. Johnson S, Healey A, Evans J, Murphy M, Crawshaw M, Gould D. Physical and cognitive task analysis in interventional radiology. *Clin Radiol* 2006;61(1):97-103.

85. Gould DA, Reekers JA, Kessel DO, et al. Simulation devices in interventional radiology: validation pending. *J Vasc Interv Radiol* 2006;17(2 Pt 1):215-6.

86. Gould DA, Healey AE, Johnson SJ, Lewandowski WE, Kessel DO. Metrics for an interventional radiology curriculum: a case for standardisation? *Stud Health Technol Inform* 2006;119:159-64.

87. Dawson DL. Training in carotid artery stenting: do carotid simulation systems really help? *Vascular* 2006;14(5):256-63.

88. 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(3):406-11.

89. Szalay D, MacRae H, Regehr G, Reznick R. Using operative outcome to assess technical skill. *Am J Surg* 2000;180(3):234-7.

90. Maithel S, Sierra R, Korndorffer J, et al. Construct and face validity of MIST-VR, Endotower, and CELTS: are we ready for skills assessment using simulators? *Surg Endosc* 2006;20(1):104-12.

91. Grober ED, Hamstra SJ, Wanzel KR, et al. Validation of novel and objective measures of microsurgical skill: Hand-motion analysis and stereoscopic visual acuity. *Microsurgery* 2003;23(4):317-22.

92. Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. *Am J Surg* 2006;191(1):128-133.

93. Wang EE, Vozenilek JA. Addressing the systems-based practice core competency: a simulation-based curriculum. *Acad Emerg Med* 2005;12(12):1191-4.

94. Madan AK, Frantzides CT, Quiros R, Dujovny N, Tebbit C. Effects of a laparoscopic course on student interest in surgical residency. *Jsls* 2005;9(2):134-7.

95. Beard J, Strachan A, Davies H, et al. Developing an education and assessment framework for the Foundation Programme. *Med Educ* 2005;39(8):841-51.

96. Berry M, Lystig T, Beard J, Klingenstierna H, Reznick R, Lonn L. Porcine Transfer Study: Virtual Reality Simulator Training Compared with Porcine Training in Endovascular Novices. *Cardiovasc Intervent Radiol* 2007;30(3):In press.

97. Ziv A, Wolpe PR, Small SD, Glick S. Simulation-based medical education: an ethical imperative. *Acad Med* 2003;78(8):783-8.

98. Gruber FP, Hartung T. Alternatives to animal experimentation in basic research. *Altex* 2004;21 Suppl 1:3-31.

99. Fitzpatrick A. Ethics and animal research. *J Lab Clin Med* 2003;141(2):89-90.

100. Machan TR. Putting humans first : why we are nature's favorite. Lanham, Md.: Rowman & Littlefield Publishers, 2004.

101. Schijven M, Jakimowicz J. Virtual reality surgical laparoscopic simulators. *Surg Endosc* 2003;17(12):1943-50.