Effectiveness of Virtual Reality Based Immersive Training for Education of Health Professionals: a Systematic Review

Thesis submitted in partial fulfilment of the requirements for the Degree of Master of Health Sciences in the University of Canterbury by Prasanna Karunasekera

University of Canterbury

2011
# Table of Contents

Acknowledgements ..................................................................................................................... 4

Abstract ....................................................................................................................................... 5

List of tables ................................................................................................................................ 7

List of figures ................................................................................................................................ 8

Objectives of the study .................................................................................................................. 9

Introduction ................................................................................................................................... 10

  Problem with clinical education ................................................................................................. 13
  Training on animal models ........................................................................................................... 15
  Simulation-based medical education ............................................................................................ 16
  Virtual Reality (VR) simulations .................................................................................................. 19
  VR simulations – background and history .................................................................................. 23
  VR simulation in health professional’s education ......................................................................... 23
  Medical education research on VR training ............................................................................... 27

Methods ........................................................................................................................................ 29

  Framing the research question ................................................................................................... 30
  Search of the literature ............................................................................................................... 31
  Assessment of study eligibility .................................................................................................. 33
  Appraisal of included studies ..................................................................................................... 34

Results .......................................................................................................................................... 40

  Summary results from included studies .................................................................................. 42
  Types of simulators ...................................................................................................................... 78
  Emergent themes from the studies appraised ............................................................................ 84
Discussion........................................................................................................................................... 91

Conclusion................................................................................................................................................100

References.................................................................................................................................................. 101

Glossary......................................................................................................................................................107

Appendix A: Included studies.................................................................................................................. 112

Appendix B: Excluded studies annotated by reason for exclusion......................................................... 115

Appendix C: Summary of individual study appraisals.............................................................................211
Acknowledgements

I would like to express my deep and sincere gratitude to my supervisor, Dr. Arindam Basu, Senior Lecturer, Health Sciences Centre, University of Canterbury, whose supervision, advice, and guidance from the very early stage of this research as well as giving me extraordinary experiences throughout the work. I appreciate his vast knowledge and skills in many areas, and his detailed and constructive comments. His understanding, encouraging and personal guidance have provided a good basis for the present thesis.

I wish to express my warm and sincere thanks to Associate Professor Ray Kirk, Director, Health Sciences Centre, University of Canterbury, for his important guidance and support throughout this work.

This project was funded by the Tertiary Education Commission funded research project on Virtual Reality and immersive training for medical education, currently being conducted at the Health Sciences Centre, University of Canterbury.
Abstract

Virtual Reality (VR) refers to computer generated artificial environment in which one’s actions partially determine what happens in the environment. In medical education and training, VR simulators use computer-generated objects on computer interface and allow the trainee or student to manipulate objects to receive feedback on the performance. The purpose of this thesis is to synthesize evidence on the effectiveness of different virtual reality based immersive training tools for health professionals.

A systematic review of the literature was conducted to assess the effectiveness of VR tools in the training of health professionals. A focused search of literature resulted in an initial retrieval of 1379 relevant titles and abstracts of peer reviewed publications. All retrieved articles were initially evaluated based on titles and abstracts to identify studies to be retained for further analysis based on full text appraisal, using the Participant-Intervention-Comparator-Outcome (PICO) criteria. After review of titles and abstract, a total of 24 publications were selected for final review. Of the 24 studies identified as eligible, one was a meta-analysis, another was a systematic review, two were other types of reviews, ten were reports of Randomised Control Trials and eleven were observational or quasi-experimental studies.

Critical appraisal of these studies resulted in identification of 12 different types of applications and nine overlapping ‘themes’ related to VR simulations and clinical skill training. In general, findings from this review indicate that, VR simulators can be considered a useful tool for improvement of clinical skills performance especially for novices with limited experience.
In combination with existing opportunities to work with real patients, VR based training can increase the range of experience to learn about and deal with medical problems as learners and practitioners. The current evidence on the effectiveness of using VR training applications for improvement of clinical skills of health professionals is limited but sufficiently encouraging to justify additional clinical trials in this area.

There remain several limitations in the research on the effectiveness of using VR training environments for health professionals. Further research work is required on role of virtual reality simulators in the transfer of skills, optimal VR applications and their technology for each discipline in medicine, economical success and other issues of VR usage, if medical skills laboratories are to remain an integral component of medical education.
List of tables

Table 1: The Six Main Categories of Bloom's Taxonomy .............................................. 11

Table 2: Included studies: brief characteristics and main findings ............................... 59

Table 3: List of the VR applications used for the training of health professionals .......... 72

Table 4: Key themes identified as a result of appraisal of articles ............................... 79
List of Figures

Figure 1: LAP Mentor multi-disciplinary LAP surgery simulator (Simbionix USA Corporation) ................................................................. 22

Figure 2: Application of selection criteria to citations ................................................. 41

Figure 3: Sub-sets of the 24 included studies ................................................................ 43
Objectives of the thesis

The objectives of this thesis were as follows:

1. Conduct a systematic review of the literature to assess the effectiveness of Virtual Reality based immersive training applications for improvement of clinical expertise of health professionals

2. Identify and list effective Virtual Reality based simulation tools for the training of health professionals
Introduction

Medical education has changed over the years and affected in different ways including: changing demographics of patients and professionals; emergence of new diseases; technological innovations; consumer empowerment; increased involvement and self organization of patients in clinical decision making, and emphasis of effectiveness and efficacy in health care decisions. As a result, at present, there are significant societal demands and budgetary restrictions to improve the quality of medical education and the safety of medical care (Vozenilek, et al., 2004).

In order to teach most effectively, educators must have clear objectives of the skills we wish students to master after having received the benefit of teaching. In deciding which activities to include in educational software, Bloom’s Taxonomy of Educational Objectives (Bloom and Krathwohl, 1956) is widely used by educators today to judge the depth and appropriateness of their coverage of course material. As discussed by Bell et al. (1995), the general categories of Bloom's Taxonomy are shown in table 1, and range from knowledge (memorization) at the low end, to evaluation (judgment) at the high end.
### Table 1: The Six Main Categories of Bloom's Taxonomy (Bell et al., 1995)

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1, Knowledge</td>
<td>The basic ability to recall information, without requiring any understanding of the material being recalled.</td>
</tr>
<tr>
<td>Level 2, Comprehension</td>
<td>The ability to understand and interpret material or situations, and to extrapolate that understanding to areas not covered by the original input</td>
</tr>
<tr>
<td>Level 3, Application</td>
<td>The ability to determine which knowledge is relevant to a particular situation, and to correctly apply that knowledge to produce a correct solution to the problem at hand.</td>
</tr>
<tr>
<td>Level 4, Analysis</td>
<td>The ability to break a complex problem or situation into parts, and to recognize the relationships between the parts and the organization of the parts.</td>
</tr>
<tr>
<td>Level 5, Synthesis</td>
<td>The ability to create a unique new entity, by drawing on different aspects of knowledge and understanding, such that the result is more than simply the sum of its component parts.</td>
</tr>
<tr>
<td>Level 6, Evaluation</td>
<td>The ability to judge the value of ideas, solutions, methods, etc. This level is considered to be the top of the cognitive hierarchy.</td>
</tr>
</tbody>
</table>

The higher levels of Bloom's taxonomy begin with analysis and include synthesis and evaluation. Bloom ranked evaluation as the highest level in his taxonomy on the grounds that a thorough evaluation requires the use of all five of the lower levels (Bell et al., 1995). When developing instructional software, it is important to keep in mind the specific skill levels and learning styles which one is attempting to address, to ensure that the software is fulfilling a need which could not be met using simpler, less costly methods (Bell et al. (1995).
The authors pointed out that students learn through a variety of different mechanisms, many of which are not utilized adequately in traditional educational methods. In addition to not addressing a particular students "preferred" learning style, authors commented that students learn more and retain more when information is presented to us multiple times, preferably through multiple channels.

The recent technological advances have made available to health-care professionals a wide set of innovative Virtual Reality (VR) training tools. Virtual reality stands poised to not only add to the variety of educational delivery mechanisms, but to specifically address those areas where traditional methods are weakest. Bell and Fogler (1997) tested the efficacy of VR based computer modules for use in undergraduate engineering education. The authors commented that traditional teaching and testing methods tend to stress primarily the three lowest levels of Bloom's taxonomy -- knowledge, comprehension, and application, whereas VR provides an environment in which students can exercise the higher levels of Bloom's taxonomy such as level 4, analysis, level 5, synthesis, and level 6, evaluation, in a manner totally unique from other educational methods. They also pointed out that within a VR simulation, students are free to explore, and to examine their environment from any viewpoint they desire including hazardous and inaccessible locations. This enable users to move beyond "real-world" experiences by interacting with or altering virtual objects in ways that would otherwise be difficult or impossible. Authors believed that with this newfound freedom to explore, students can analyze their problems and evaluate possible alternatives in ways never before possible. It stands poised to not only add to the variety of educational delivery mechanisms, but to specifically address those areas where traditional methods are weakest. The rest of this article examines the evidence about the effectiveness of VR based immersive training aimed at improvement of clinical expertise of health professionals.
**Problem with clinical education**

The field of medicine, physicians in general learn by initially observing, then practicing, under the supervision of a more experienced physician. The traditional system of clinical education relies on practicing diagnostic, therapeutic, and procedural skills on live patients. The challenge of clinical teaching is how to balance patient care responsibilities with teaching opportunities. Clinical teaching occurs in fast-paced and chaotic surroundings where simultaneous – and often competing – demands are placed on all members of the health care team. It has been much criticized for its variability, lack of intellectual challenge, and haphazard nature (Spencer, 2003). Hourly work limitations are placing a burden on the training of physicians. Added to time pressure is the challenge of providing instructions to learners at different development levels. Better patient care and optimal physician training are often mutually exclusive in the clinical setting, and consequently live-patient training has several significant shortcomings.

Reznek and colleagues (2002) argue that:

“Clinical education is not ideal as clinical practice of medicine has been refined over the years specifically to improve patient care and not necessarily education”. In addition, this training system is also inefficient in that the trainee is not even guaranteed the opportunity to learn a procedure” (Reznek et al., p.78-87). Clinical education as traditionally practiced provides inadequate opportunity to the trainee to learn the procedures adequately and that there is a disjoint between medical education and medical practice if medical education is only imparted in the setting of a clinical practice. Hewison and Wildman (1996) pointed out that “a clinical environment is meant for care delivery not learning”. In their research, they identified existence of a theory-practice gap in nursing in the United Kingdom as issue of concern for many years.
The concept of ‘‘learning by doing’’ has become less acceptable, particularly when invasive procedures and high-risk care are required (Vozenilek et al. 2004). From the viewpoint of the instructor, allowing a novice to perform a skill on a patient who may develop serious complications or even may die from the technical inadequacy is untenable from an ethical or legal perspective. Even more disturbing to the learning process is the fact that an instructor is ethically bound to stop the resident if he or she is making an error. For this reason, the resident will only rarely have the opportunity to experience complications resulting from his or her actions.

Furthermore, problems occur in clinical training, or in areas where students must learn complex processes, but have limited access to real world work spaces, or need to get greater understanding of processes that cannot be seen in the real world, such as chemical reactions, or audio transmissions. There is no guarantee that clinical placements will provide appropriate learning opportunities for students to experience low incidence but high risk clinical events that prepare them for safe and effective clinical practice (Murray et al., 2008). Finally, live-patient dependent education is not cost efficient. Haluck (1999, as cited in Reznek et al., 2002) pointed out that ‘‘live’’ human being treated as a model can be very expensive both in terms of time and money. The instruments are not reusable, and it takes longer for a trainee to perform the procedure. In addition, an attending observer ideally should be present at all times, keeping him or her from other clinical responsibilities.

These limitations decrease the patient experience that a resident experiences, thus limiting exposure to clinical situations and procedural practice opportunities. It is obvious that the live patient ‘‘model’’ is not an ideal instrument for education, especially for the introductory instruction of procedures and most medical management procedures.
This form of apprenticeship is increasingly becoming disfavoured by an ever more safety-conscious public. As a multitude of factors play increasingly important roles in the execution of the medical knowledge acquired during one’s education, it is essential to improve the training programme by supplementing the standard teaching methods to increase the number of learning modes available to the student. Given the complexity, conflicting roles and tensions of the clinical setting, the challenge facing clinician-teachers is to create a high quality learning environment for students.

*Training on animal models*

Research laboratories using animal models (especially pigs) are very popular for exploring the safety and feasibility of new surgical procedures, instruments, medications or materials. They also provide a realistic safe environment where randomised controlled studies could be carried out before starting human trials. Training on cadaver specimens enables the trainee to anatomy identical to that found in the living. Human cadaver surgical skills practice has been found to impart greater confidence and skills mastery to training course participants (Blaschko et al. 2007).

Limitations on widespread use of human anatomical material for surgical skills training include their cost and limited supply, difference in quality of tissues found in these specimens and the fact that cadavers may traditionally only be used once. Cadavers are expensive, involving initial transportation and preparation, large space for storage and finally disposal, burial or cremation. Blaschko et al. (2007) point out that the use of animals in training health-care personnel is being banned in countries like Britain and Canada due to the concern of contamination of health-care workers with bovine spongiform encephalopathy (BSE) virus and other livestock-infecting diseases.
Hoffman (1997, as cited in Ahlberg et al., 2002) pointed out that “in recent years, the tradition of training surgical students in animal models and human cadavers come under fire for ethical reasons”. Animals and cadavers were popular in surgical training until box trainers were developed, encompassing laparoscopic video equipment.

*Simulation-based medical education*

Simulations are abstractions of reality. Often they deliberately emphasize one part of reality at the expense of other parts. It is a process for practice and learning that can be applied to many different disciplines and trainees including aviation, nuclear power plants, space aeronautics, the military, business, and healthcare. Simulation facilitates learning through immersion, reflection, and practice - minus the risks inherent in a similar real life experience. In medical simulation, computer-controlled equipment advances medical learning and ensures that students learn procedures and treatment protocols before using them on actual patients. Simulation based medical education provides a controlled environment that imitates a real-life patient care setting and allows students and providers to learn, practice, and repeat procedures as often as necessary in order to correct mistakes, fine-tune their skills, and optimize clinical outcomes (Murphy et al., 2007). In addition, with simulation, students and residents can gain experience with various types of patients and cases they may not actually encounter during their rotations and shifts.

The systematic review by Issenberg et al. (2005) examined the use and effectiveness of simulation technology in medical education. Specifically they addressed the following question: “what are the features and uses of high-fidelity medical simulations that lead to most effective learning?” The review stated that simulations are now in widespread use in medical education and medical personnel training to boost the growth of learner knowledge.
and safe practice opportunities. The reviewers concluded that “while research in this field needs improvement in terms of rigor and quality, high-fidelity medical simulations are educationally effective and simulation-based education complements medical education in patient care settings”. Through simulation, learners can practice and master skills while protecting patients from unnecessary risk. Ziv et al. (2003) in their article of simulation based medical education reported that use of simulation-based learning improves health professionals' knowledge, skills, and attitudes while enhancing patient safety. They also showed that advanced simulation-based learning can provide realistic representations of complex clinical environments and allow educators to alter patient reactions and responses in ways unattainable with actual patients.

Participants of simulation are immersed in these imitations of real-life experience and play roles in scenarios such as cardiac resuscitation teams, procedural performance, delivery of babies, providing anaesthesia, surgical operations, dentistry, and nursing care, to name just a few. The use of simulation spans a spectrum of sophistication, from the simple reproduction of isolated body parts through to complex human interactions portrayed by simulated patients or high-fidelity human patient simulators replicating whole body appearance and variable physiological parameters.

Ziv et al. (2003) point out that simulation has been used unsystematically since the early days of medicine (p. 783-788). They reported that in the 16th century, mannequins (referred to as ‘phantoms’) were developed to teach obstetrical skills and reduce high maternal and infant mortality rates. Modern mannequin-based simulator has a computer representation of the patient similar to that in a desktop simulator, replacing the videos, drawings, and animations with actual functions of the “plastic person.” Full body mannequin simulators originated in the field of anaesthesia.
Denson and Abrahamson (1969) first introduced human patient simulation to the medical community in 1969, when they used a patient anaesthesia simulator to augment resident training. This model, known as Sim One, simulator consisted of a life-size mannequin connected to a computer, an instructor’s console, an interfacing unit, and an anaesthesia machine.

Fritz et al. (2008) reviewed the published literature of mannequin-based simulation in emergency medicine focusing on high-fidelity simulation (HFS), in particular the advantages and disadvantages of using this technology. The review showed that simulation has many potential advantages over traditional learning methods currently employed in training medical, nursing and paramedical staff (Fritz et al., 2008). These include: no direct risk to patients, potential to increase the speed of acquisition of clinical skills, allow a standardized curriculum to be developed, team training and continuing education.

The main disadvantages to mannequin-based to simulation are the significant cost involved and the potential for negative transfer or learning. Murphy et al. (2007) point out that “mannequins reproduce many of the features of critical life threatening illness but often have technical and hardware limitations falling short of ideal for optimal simulation of clinical reality”. Bond et al. (2007) reported put that “medical simulation is a rapidly expanding area within medical education” (p. 353-63). They showed that mannequin-based simulator has been used particularly in emergency medicine over a decade for various purposes from teaching procedures to team-based training. Newest full body simulator models incorporate computerized models that approximate the physiology seen in the human body. These models have uses beyond anaesthesia and are now also used for surgery, critical care, obstetric, emergency medicine, and internal medicine (Gaba, 1999, as cited in Dianesliwka, 2008).
Last decade has seen an explosion of the number of tools available to enhance medical education: web-based education, virtual reality and high-fidelity human patient simulation (Vozenilek et al., 2004). Reznek et al. (2002) pointed out that technologic advances in the areas of virtual reality and computer enhanced simulation have introduced a new method of teaching that bypasses each of the ethical, financial, and practical deficiencies of live patient training that have been illustrated in the following section.

The traditional system of clinical education relies on practicing diagnostic, therapeutic, and procedural skills on live patients. At the same time, there is an obligation to provide optimal treatment and to ensure patients' safety and well-being. Ziv et al. (2003) point out that these conflicting needs generate a fundamental ethical tension in medical education. They found that simulation-based learning can help alleviate this tension by developing health professionals' knowledge, skills, and attitudes while protecting patients from unnecessary risk.

**Virtual Reality simulations**

VR is a human-computer interface that simulates a realistic environment and allows users to interact with it. Ertan defined VR as follows: “a collection of technologies that allow people to interact efficiently with three dimensional (3D) computerized databases in real time, using their natural senses and skills” (Ertan, 2004, 23:147–152). Virtual reality is essentially a combination of computer hardware architecture and software programming designed to immerse users in artificially-created virtual environments such that users perceive themselves to be included in and interacting in real-time with the environment and its contents. Virtual reality implementations typically use high speed, high quality three dimensional graphics,
and 3-D audio and specialized hardware such as head-mounted displays and wired clothing to achieve high degrees of realism and believability. Eslinger (1993, as cited in Onyesolu & Eze 2009) suggest that this simulated alternative form of education facilitates learning due to the ability of human brain to comprehend images much faster than they grasp lines of text or columns of numbers. With this newfound freedom to explore, students can analyze their problems and evaluate possible alternatives in ways never before possible.

Virtual Reality (VR) provides opportunity for enhancing and modifying the learning experience. Mantovani et al. (2003) identified virtual environments (VEs) as an attractive educational context which can provide a rich, interactive and supporting experiential learning. A virtual environment presents a unified workspace allowing all the functions to be located in the same physical space. Most current virtual reality environments are primarily visual experiences, displayed either on a computer screen or through special stereoscopic displays, but some simulations include additional sensory information, such as sound through speakers or headphones. Some advanced, haptic systems now include tactile information, generally known as force feedback, in medical and gaming applications. A virtual world allows human participants to perform tasks as ‘naturally’ as they would do in everyday reality (Usoh and Slater, 1995). While the use of simulation in non digital forms has been used in health professional education for many years, recent improvements in technology have created highly realistic simulators capable of very high levels of fidelity, to the point that making the situation appear to be quite real.

Medical and surgical patient care is one of the major application areas for virtual reality (Riva, 2003). Virtual reality offers promising solutions in many areas of medical care including remote and local surgery, surgery planning, medical education and training,
treatment of phobias and other causes of psychological distress, skill training, and pain reduction. Education and training is one of the most promising application areas for virtual reality technologies (Riva, 2003). Computerised three dimensional atlases presenting different aspects of the anatomy, physiology, and pathology as a unified teaching atlas whereas systems based on virtual reality offer a unique opportunity for the training of professional surgical skills on a wide scale. Virtual reality programs have also been used for a variety of medical emergency, mass casualty, and disaster response training sessions for medical and public health professionals. More recently, the scope of VR applications in medicine has broadened to include physical and psychiatric rehabilitation and, to a lesser extent, diagnosis.

In the past decade we are seeing the emergence of medical applications for virtual reality technologies (Székely & Satava, 1999). These include telepresence surgery, three-dimensional (3-D) visualization of anatomy for medical education, VR surgical simulators, and virtual prototyping of surgical equipment and operating rooms. These applications present dynamic, three-dimensional views of structures and their spatial relationships, enabling users to move beyond "real-world" experiences by interacting with or altering virtual objects in ways that would otherwise be difficult or impossible. These new systems make broad-based training experiences available for trainees at all levels and trainees can acquire proficiency and gain confidence in the ability to perform a wide variety of techniques long before they need to use them clinically.
The LAP Mentor multi-disciplinary LAP surgery simulator enables simultaneous hands-on practice for a single trainee or a team. The system offers training opportunities to new and experienced surgeons for everything from perfecting basic laparoscopic skills to performing complete laparoscopic surgical procedures.
Virtual Reality simulation - background and history

Uranus (2004, as cited in Aggarwal et al., 2006) pointed out that use of VR simulators was inspired by the successful use of such tools to train airline pilots. They realised that simulators have the ability to create a ‘realistic, real-time’ environment, with inbuilt objective assessment of performance, offering a feasible alternative to live training where pilots can train in a safe environment accelerating their learning curve without risking their or their passengers’ lives. First introduced to surgery in 1991, acceptance of VR training has been slow, partly because of scepticism within the medical community but also due to the lack of well-controlled clinical trials (Riva, 2003).

Since 1991, increasingly sophisticated developments have occurred worldwide to improve the learning of individual and team reasoning, communication, and technical and other skills through the development of medical skills laboratories that employ various levels of simulation (Maran, et al 2003; Gorman, et al 2000). Factors associated with adoption and popularity of using simulation in medical learning environment include growth in medical knowledge base, increase in accountability outcome measurement, and changes in the pattern of medical education.

Virtual Reality (VR) simulation in health professional’s education

Employed most commonly for examination, surgical, and endoscopic procedures training and assessment, Virtual Reality (VR) simulation has been offered as a solution to improve medical training. Zajtchuk & Satava (1997) point out that “field of medical education was the first medical discipline to take advantage of the power of virtual environments.” In the recent years, there has been growing interest in the use of virtual reality simulators as adjuncts to conventional training (p. 63-64).
VR simulators have some clear benefit over number of shortcomings of other methods of instruction in medical education, not only in terms of teaching operative skills, but also in terms of assessing those skills. Given the environment of decreasing clinical exposure and pressures for improved learner outcome measurements, simulation offers suitable alternative to allow student learning and initial expressions of competence to take place in a patient-free environment.

Several pioneer research groups have already demonstrated improved clinical performance using VR imaging, planning and control techniques. For instance, Stevens (2005, as cited in Persky & McBride, 2009) reported that virtual environment technology interactive simulated patient training tools have proven effective in increasing health professionals’ knowledge and competencies in safe learning environments where failure does not hold the grave consequences of real clinical experience. Although, evidence for its use in patient safety might still be under-researched it has the ability to providing a non-threatening environment in which trainees, not yet achieving proficiency, may practice a skill with the freedom to fail, without involving unpleasant consequences, or wasting consumable training materials (Aggarwal and Oliver, 2010).

Hamilton (2002, as cited in Aggarwal & Moorthy, 2004) randomized surgical trainees to ten half-hour sessions on a box trainer or to the MIST-VR (Minimally Invasive Surgery Training-Virtual Reality), with baseline and post-training skill assessments on both trainers. All achieved significant improvements regardless of which simulator they had trained on, although the magnitude of improvement for the MIST-VR trained group was significantly greater (P< 0.05) than that for the box-trained group on both forms of assessment. This study was one of the first to demonstrate improvement of psychomotor skills and intra-operative
performance after training on VR systems, underscoring the importance of such systems for the training of surgeons.

Simulation-based medical education (SBME) can be a valuable tool in create cases involving uncommon clinical situations and allow training to perform risky procedures or manage difficult clinical situations without patient involvement. Persky & McBride (2009) point out that “simulation-based training allows presenting and/or manipulating characteristics and objects that in real life are irreversible, invisible, or insubstantial” (p. 677-682). Highly developed SBME allow educators to alter patient reactions and responses in ways unattainable with actual patients (Ziv et al. 2003).

Another potential benefit of virtual reality simulation is the ability to repeat training procedures. It enables trainees to practice repeatedly until their skills are mastered and to maintain those skills once acquired. Because simulators are potentially available at all hours, training in operative technique may be tailored around work and other commitments and incorporated into a programme of instruction designed to facilitate progression from novice to expert. The trainee can also be mentored through the case, stopping as necessary to explain difficult parts of the procedure.

Krummel (1998, as cited in Aggarwal & Black, 2006) pointed out that benefits of VR simulation are the possibility to train in an educationally-orientated environment free of the time to develop new technology and modules, repeatability, overall lower cost of material development. By its very nature, Virtual Reality simulation depends on tracking measurements of instrument movement which may easily be extracted from the simulator platform to provide objective measurements of the user’s performance in executing the simulated task.
Objective comparison of trainees has not been possible in the past, but providing consistent replication of patient cases and impartial conditions generated by virtual reality simulators represent a positive step in this direction, allowing the skills of one trainee to be directly compared with those of another. This represents the first time in the history of surgical training that entirely objective assessments of psychomotor skills have been possible (Cosman et al, 2002). They also identified that, these simulators can offer performance assessment without the need for monitored human supervision, and directly measure multiple aspects of a subject’s psychomotor performance on specific skills. The review by Aggarwal and Moorthy (2004) illustrated that VR based training facilitates opportunity to provide formative and summative feedback regarding technical skills. The authors stated that VR training offer an objective assessment of performance. Aggarwal and Hance (2004) noted that this data can then be used to develop a training programme which is completed upon the demonstration of pre-defined levels of proficiency.

Simulation-based training can also process real patient data from a CT scan, and enable the trainee to practice the ‘real’ case on the simulator, prior to performing the real case on the patient (Aggarwal & Black, 2006). Any tricky or difficult parts of the procedure can be repeated, reducing the likelihood of real errors or adverse events occurring due to technical difficulties. By registering images obtained from real patients, the simulated tasks may be varied in order to avoid learning of the simulation, and to expose the trainee to a variety of anatomical variations under various physiological conditions.

In general, the review of papers on benefit of VR simulation shows that VR can be considered a useful tool for education and training in medicine. When the time comes to transfer these skills to the real situations, Virtual Reality simulators are likely to be
successful. In the review literature, few studies have assessed the transfer of skill using VR simulation to actual operation room performance and showed improved trainee performances. In their review Grantcharov et al. (2004) reported that VR trained skills may be transferable and lead to improved performance intra-operatively compared to standard training.

*Medical education research on Virtual-reality training*

There is a growing body of evidence that simulation technology provides an effective mechanism to educate and evaluate health professionals. Concurrent with the expansion of simulation technology in medical education is a growing call for higher quality in medical education research. To date, there has been only little critical evidence based in the form of a systematic review of either randomized trials or other forms of intervention studies. The first of these, by the Australian Safety and Efficacy Register of New Interventional Procedures–Surgical (ASERNIP-S) in 2003, focused on the effectiveness of surgical simulators only, and identified 26 randomised, controlled trials (RCTs) that met the review inclusion criteria (Sutherland L.M., et al). Findings from the review showed that, compared with no training at all, computer simulator training generally showed greater improvements in surgical skills but was 'not convincingly superior' to video simulator or standard (surgical drills) training. While there may be compelling reasons to reduce dependence on patients, cadavers, and animals for surgical training, none of the methods of simulated training (including computer simulation) have not been shown to be better than other forms of surgical training. Aim of the more recent review (Marita Lynagh et al., 2007) was to analyse systematically existing research evidence on the effectiveness of medical skills laboratories, with a particular focus on issues of transferability to clinical practice and retention of skills over time.
A range of databases was utilized to search for relevant papers published from 1998 to June 2006. A total of 44 RCTs were identified for inclusion in the review. In their conclusion, the reviewers showed that medical skills laboratories do lead to improvement in procedural skills compared with standard or no training at all when assessed by simulator performance and immediately post-training. They commented on the lack of well designed trials addressing the crucial issues of transferability to clinical practice and retention of skills over time.

Despite the fact that several single descriptive and some analytical studies on the effectiveness of virtual reality medical training systems on the learning of residents and trainee physicians (primarily endoscopists and surgeons) are available, there are no meta-analyses or integrative reviews that describe the overall effectiveness of this approach. While not yet in the mainstream of academic medical training, many VR applications are emerging, with target audiences ranging from first- and second-year medical students to residents in advanced clinical training. It is therefore timely to examine what is the overall effectiveness of a successful stimulation in terms of learning and clinical outcomes.

The purpose of this thesis is to provide an overview and production of the evidence pertaining to the effectiveness of Virtual Reality (VR) based immersive training applications for improvement of clinical expertise of health professionals. This will be done by conducting a systematic review of the literature on the effectiveness of Virtual Reality as a tool for the training of nurses and physicians in training and other health professionals. The review examines the evidence based on the overall effectiveness of VR simulation in terms of learning and outcomes and in addition, identifies VR based simulation tools for the training of health professionals. It is expected that this review will provide information that will be of direct interest to the medical educators.
Method

The goal of this thesis was to characterize the effectiveness or efficacy of virtual reality based immersive training application for health professionals. This was addressed by addressing the following two objectives: First, a systematic review of the literature to assess the effectiveness of Virtual Reality based immersive training applications for improvement of clinical expertise of health professionals was conducted. Second, based on the review, effective Virtual Reality tools for health professionals were identified and listed.

The steps of this systematic review were as follows:

1. A question was framed based on the Participant-Intervention-Comparator-Outcome (PICO) criteria
2. A systematic search of the literature was conducted with specific dates, specific terms and selected databases
3. The publications were initially evaluated with respect to their titles and abstracts and then they were further appraised with respect to close reading of their full texts whenever possible judging the extent to which these publications and the methods contained in those studies had conformed to internal validity of the studies.
4. For the subset of peer reviewed, internally valid studies that passed the selection criteria of the critical appraisal, the findings were summarized
5. Finally, themes were identified from these studies and summarized here along with a list of tools of virtual reality used for this study.
Framing the research question

The Centre for Clinical Effectiveness (CCE) at Monash University developed a framework called “PICO” (Population-Intervention-Comparator-Outcome) to make the process of asking an answerable question easier and to develop and refine search approach. The key research question for this review was formulated using the Patient-Intervention-Comparator-Outcome (PICO) framework. The framework is briefly explained below.

**PICO** stands for:

**P:** Population
- What are the characteristics of the population?

**I:** Intervention or exposure
- What do you want to do with this patient/population (e.g. treat, diagnose, observe)?

**C:** Comparison or control
- What is the alternative to the intervention (e.g. placebo, different drug, surgery)?

**O:** Outcome
- What are the relevant outcomes (e.g. morbidity, death, complications)?

PICO is a method of putting together a search strategy that allows taking a more evidence based approach to literature searching. It allows identifying the information needed to answer the question and to translate the question into searchable terms.
In this context, the research question was framed as follows:

“Compared to usual or traditional approaches, for physicians, nurses, and other health professionals directly involved in patient care in a number of care settings, what is the overall effectiveness of virtual reality based training programmes”?

Search of the literature

A systematic method of literature searching and selection was employed this review. Searchers were limited to English language material published from 2000 -2009 in view of the advancement of technology. Using the specified search terms, a number of bibliographic databases were searched to identify an initial set of relevant reports of studies.

The following databases were used to identify key literature:


2. CINAHL (Cumulative Index of Nursing and Allied Health Literature) [http://www.ebscohost.com/cinahl](http://www.ebscohost.com/cinahl)

3. Eric (Educational Resources Information Centre) [http://www.eric.ed.gov](http://www.eric.ed.gov)

4. Cochrane Central Register of Controlled Trials (CENTRAL) [http://www.thecochranelibrary.com](http://www.thecochranelibrary.com)

5. Cochrane Database of Systematic Reviews [http://www2.cochrane.org/reviews](http://www2.cochrane.org/reviews)

6. Embase [www.embase.com](http://www.embase.com)

7. Health Technology Assessment Database
8. Google search [http://scholar.google.co.nz](http://scholar.google.co.nz)

Hand searching of journals, contacting of manufacturers, or contacting of authors for unpublished research was not undertaken in this review. The following search terms were used to search the databases:

1. For searching systematic reviews of the relevant studies, the Montori Systematic Review search filter was used (Montori et al. 2005)

2. For the other types of studies, the following search terms were used in different combination using Boolean operators:

   - “virtual reality”
   - “medical education”
   - “medical training”
   - “virtual world”
   - “physician education”
   - “improvement in medical knowledge”
   - “nursing education”
   - “virtual stimulator”
   - “web based training”
   - “residency”
   - “medical diagnosis”
   - “operative skills”
   - “surgical procedure”
Using the specified search terms, bibliographic databases were mined to identify current research on effectiveness of virtual reality based immersive training applications for improvement of clinical expertise of health professionals. Systematic reviews were considered in this thesis for following reasons. Firstly, the purpose was to identify the best evidence of what works. According to the evidence hierarchy (Evan’s 2003), a systematic review belongs to category 1 of good evidence. Secondly, this thesis review evaluates a technology; the technology being tools based on virtual reality and their effectiveness in teaching specific skills. Established practices of technology review mandate that not only primary studies but also secondary data sources should be reviewed for generation of evidence (NHMRC, 1999).

Assessment of study eligibility

An article was included for further consideration if it meets one of the following criteria:

- Published between 2000 – 2009

- Published in English language or appropriate translations were already available

If an abstract was not available despite all possible means of search and retrieval, then the resource was not considered for further review. If an article was not available as a separate publication (i.e. as a primary study or a secondary data analysis) then the study was excluded. Thus, if a study was available only as a conference presentations and but not as a separate study published in a peer reviewed journal, then that study was removed.
Therefore we followed the principals of health technology assessment and included systematic reviews. Lastly, to increase the comprehensive coverage of different studies systematic reviews were included in the search process and data extraction.

Appraisal of included studies

This study was conducted using the well established methodologies outlined in the Cochrane Reviewer’s Handbook (Higgins 2009), and using the critical appraisal criteria used by the National Health and Medical Research Council of Australia. PICO (Population, Interventions, Comparator and Outcomes) components of the question, with the additional specification of types of study that were included, form the basis of the pre-specified eligibility criteria for the review. Each article or literature based resource were critically appraised using the standard criteria for analysis of an article and extraction of data.

Appraisals of the literature were conducted in two phases:

In phase I, all retrieve articles were initially evaluated based on titles and abstracts as a first pass to identify studies to be retained for further analysis based on full text appraisal, using the PICO criteria mentioned above. The abstract was first used to identify the research question. (In the first pass stage, the theme of the article and its applicability to the research question were assessed). Abstract was then used to identify the participants for the study, the technology used, the comparison group or the comparison profile used for the study, the specific learning outcomes for the participants in the study and the extent to which the study would otherwise fulfil the conditions of a robust internally valid study.

Based on the inclusion and exclusion criteria, and only on the basis of the content of their titles and abstracts, these articles were then either retained for further analysis and appraisal
in stage two or removed from the database and indicated as such. An article was excluded for further consideration to be included in the database if it fails to meet one or more of the following criteria:

1. **Inappropriate publication:**
   
   If the publication is not a primary research, then that study was not included. By primary research we meant any research where the researchers developed a simulation device using virtual reality or virtual world based application and evaluated OR the researchers did not systematically search for and compile a list of such devices (secondary data analyses or systematic reviews).

2. **Inappropriate technology:**

   If the simulation device developed by the researchers was not based on virtual reality or virtual world, then that article was removed from the review.

   By virtual reality/virtual world, we meant that the simulation may have a component that is based on either haptic or non-computer based technology, but in addition, they must have a component where computer programmes that run virtual reality or augmented reality programmes using for the training. For example, a simulation programme based on computer simulation of a surgical scenario was included while; a simulation based only on video devices and simulation hardware (for example non haptic training material) was excluded.

3. **Inappropriate participant profile:**

   We only accepted studies involving health care professionals.
Here, if the trainee population consisted of medical/nursing/allied health care students, but not qualified professionals, then that study was not taken into account. So, studies on surgical residents and interns were accepted (even if they are in addition to any medical students). Even though the review becomes limited in scope; the justification for excluding studies carried out with medical student participants is presumably because the objective was to identify VR tools that are effective for improving clinical skills in health professionals. Medical students are not in a position to make independent decisions at least in the contexts where they work, and therefore synthesis of evaluations that included medical students would not add anything to the intended audience (i.e. post certification health professionals). Thus the pedagogical inference becomes somewhat limited and review findings would only be applied in the context of post certification health professionals.

4. Inappropriate/ absence of outcomes:
   If the study only described a technology but did not describe an evaluation of the technology, then that study was not accepted for inclusion.

5. Inappropriate or absence of comparison groups:
   Because we are only going to accept papers where the authors have described a technology and in addition, they have described an evaluation of this technology to reflect how well it works, therefore it is implied that the authors may have included some sort of comparison group (either the participants are their own comparators, or the participants are compared with other participants groups etc). If no such comparisons were available, then we excluded that study in the final report based on
the full text appraisal. In this category, we did not accept studies that are just case
series unless they include some sort of a comparison usually called pre post or serial
measurements learned skills.

For example, if a technology (‘T’) was developed and tested on a sample of trainees
(‘S’), then to be included on the basis of full text appraisal, the study need to describe
at least:

a) T must have been based on virtual reality at some point, not just video
   implementations).

b) S would necessary consist of individuals who are qualified to practice
   medicine/surgery/health care delivery (interns, residents, post graduate trainees,
   physicians, surgeons, nurses are appropriate populations).

c) On the trainee sample S, there must be at least two set of evaluations. i.e. at the
   baseline start of the technology and at the end.

6. Inappropriate outcome:

   If the study outcomes were not related to learning or training of health care
   professionals, then that study was not be included in the review.

In phase II, all articles retained from phase I as eligible were evaluated and appraised. Two
steps were followed in this stage. In first step, all articles were sought for their full text
versions (this could be in the form of PDFs, other formats, or even printable texts which were
obtained through any number of ways: contacting the library of UC, searching the citation
databases, or using the reference management system to retrieve the publications, or web
services such as Google Scholar, CiteULike and others).

An article were removed from further evaluation (in step II) if a full text was not available
and mentioned as such. In the second step (phase II), all retrieved articles were evaluated on
the basis of their full text content, applying the same inclusion and exclusion criteria as
described above (In depth analysis of individual studies). Thus, in this step of the research
process, full-text versions of selected publications were critically appraised with respect to
the following criteria:

1. Study Design (whether the study is an RCT, or another comparative trial, or a
   systematic review, or meta-analysis)

2. Whether the study has an appropriate participant profile (i.e. whether the study is
   about health care professionals, and not just medical students)

3. Whether the study is about an appropriate intervention

4. The type of comparator used in the study

5. A detailed appraisal of the study outcomes were conducted to note that the outcomes
   are relevant to the question that is framed (refer to the study question)

If there is a ground of rejection of a study on any of the above criteria, the first criterion on
which a study is rejected was indicated. Additionally, all other criteria on which a study may
be rejected were indicated as well. Quality appraisal for individual studies was based on the
NHMRC criteria for assessment of individual component studies.
This generated a list of selected full text studies that were relevant for the review. This set of studies was further appraised and data extracted and evidence tables were form to indicate the following:

1) List of interventions based on virtual reality applications in the setting of a clinical workplace targeted at the training of a health care professionals

2) An evidence base to show the relative effectiveness of that virtual reality based intervention for the training of the clinical professionals.

This was done in the form of a table that shows the results from the studies in the following order:

a) Results from meta-analyses and systematic reviews

b) Results from Randomized Controlled Trials

c) Results from other types of studies that are not RCTs or meta-analyses but used other types of quasi-experimental or comparative research

These evidence tables were provide the information base for summarizing the results. A narrative summary were provided for the studies included in the research.
Results

In stage I, the search strategy identified a total of 1379 publications. As described in Figure 2 (page 40), all retrieved were assessed on titles and abstracts as a first pass to identify the relevant studies. The first phase resulted in a total of 368 non-duplicate studies. All excluded studies are presented in Appendix B, annotated by reason for exclusion based on the exclusion criteria detailed above. Reasons are presented hierarchically, such that the first reason in the list that is applied is reported.

All articles retained as eligible were evaluated and appraised. Two steps were followed in this stage of the review. Firstly, all articles were sought for their full text versions. There were 178 articles available full text or PDFs. Remaining 190 articles were removed from further evaluation as full text versions were not available and mentioned as such. Secondly, all retrieved articles were evaluated on the basis of their full text content, applying the same inclusion and exclusion criteria as described above. After consideration of full text content using the study selection criteria, a set of 24 full text publications were retrieved and scrutinised in detail for possible inclusion in the review. Reasons are presented hierarchically such that the first reason in the list that applied is reported. Other cited publications (e.g. those providing background materials) are presented in the References (Appendix A).
**Figure 2: Application of selection criteria to citations**

<table>
<thead>
<tr>
<th>Selection Criteria</th>
<th>Count (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial search</td>
<td>1379 (100%)</td>
</tr>
<tr>
<td>Abstract not available</td>
<td>196 (14.2%)</td>
</tr>
<tr>
<td>Published before 1999</td>
<td>217 (15.7%)</td>
</tr>
<tr>
<td>Inappropriate participant</td>
<td>559 (40.5%)</td>
</tr>
<tr>
<td>Main topic is not related to VR</td>
<td>13 (0.9%)</td>
</tr>
<tr>
<td>Not related to education of health professionals</td>
<td>26 (1.9%)</td>
</tr>
<tr>
<td>Articles retained for review</td>
<td>368</td>
</tr>
<tr>
<td>Retrieved full text</td>
<td>178</td>
</tr>
<tr>
<td>Inappropriate study design</td>
<td>34 (19.1%)</td>
</tr>
<tr>
<td>Inappropriate participant profile</td>
<td>16 (9.0%)</td>
</tr>
<tr>
<td>Inappropriate outcome</td>
<td>81 (45.5%)</td>
</tr>
<tr>
<td>Absence of comparison groups</td>
<td>7 (3.9%)</td>
</tr>
<tr>
<td>Inappropriate technology</td>
<td>7 (3.9%)</td>
</tr>
<tr>
<td>&lt; 10 participants in each study arm</td>
<td>9 (5.1%)</td>
</tr>
<tr>
<td>Included articles</td>
<td>24</td>
</tr>
</tbody>
</table>
Summary of results from included studies

Methodological information and results extracted from included studies are presented below. More detailed information is available in Appendix C or in the original papers. Only data relevant to the current review is presented.

Figure 3 (page 42) provides a quick guide to how the 24 eligible studies are divided. Of the 24 studies identified as eligible, one was a meta-analysis, another was a systematic review, ten were Randomised Control Trials, two were other types of reviews and eleven were observational or quasi-experimental studies. Critically appraisal of the full-text versions of selected publications were undertaken according to the criteria described in methodological section.
## Figure 3: Sub-sets of the 24 included studies

<table>
<thead>
<tr>
<th>Type of Study</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meta-analysis</td>
<td>1</td>
<td>4.16%</td>
</tr>
<tr>
<td>Systematic reviews</td>
<td>1</td>
<td>4.16%</td>
</tr>
<tr>
<td>Randomised Control Trials</td>
<td>9</td>
<td>37.5%</td>
</tr>
<tr>
<td>Other types of reviews</td>
<td>2</td>
<td>8.33%</td>
</tr>
<tr>
<td>Observational or quasi-experimental studies</td>
<td>11</td>
<td>45.83%</td>
</tr>
</tbody>
</table>
The results from the studies are summarised and briefly presented in the following paragraphs. Study characteristics and main findings are described in Table 2 (page 58).

Studies are listed in the following order:

a) Results from meta-analyses and systematic reviews

b) Results from Randomized Controlled Trials

c) Results from other types of studies that are not RCTs or meta-analyses but used other types of quasi-experimental or comparative research


The meta analysis by Gurusamy et al (2008) included a methodical search of all the relevant medical, educational and computer literature databases including the grey literature, and included randomized trials that evaluated the effectiveness of VR training. A total of 2176 primary studies were identified through electronic searches and a total of 23 trials involving 622 participants were included for final assessment. Four trials compared VR and video trainer (VT) training; 12 trials compared VR and no training or standard laparoscopic training (SLT); four trials compared VR training, VT training and no training; and three trials compared different methods of VR training. Six trials compared VR training in surgical trainees with limited experience in laparoscopic surgery. One trial did not state the experience of the participants. The other trials included medical students or surgical residents without any experience in laparoscopic surgery.
In trainees without surgical experience, VR training decreased the time taken to complete a task (95 per cent c.i. –1·50 to −0·68), increased accuracy and decreased errors compared with no training. In the same participants, VR training was more accurate than video trainer (VT) training. In participants with limited laparoscopic experience, VR training resulted in a statistically significant greater reduction in operating time, lower error score and decrease unnecessary movements than standard laparoscopic training. In these participants, the composite performance score was significantly better in the VR group than the VT group. Reviewers concluded that “VR training can supplement standard laparoscopic surgical training and it is at least as effective as video training in supplementing standard laparoscopic training”.


Aggarwal et al (2004) reviews the tools available for training and assessment in laparoscopic surgery. Medline searches were performed to identify articles and further articles were obtained by manually searching the reference lists of identified papers. Overall, the findings from this systematic review suggest that training involving box trainers with either innate models or animal tissues lacks objective assessment of skill acquisition. Virtual reality simulators have the ability to teach laparoscopic psychomotor skills, and objective assessment is now possible using dexterity-based and video analysis systems. The authors of the review concluded that there is sufficient evidence for incorporation of virtual simulation into current training programmes and they emphasized that the expanding scope of this technology should be coupled with validated training programmes.
Summary of results from Randomised Controlled Studies

Botden et al. (2008) conducted RCT on participants of several European Association for Endoscopic Surgery (EAES)-approved laparoscopic skills courses (n = 45). First participants filled out a questionnaire on their opinion on laparoscopic suturing training. After a general introduction of the simulators was given, all participants were randomly and blinded divided into two equally sized groups: group A (n = 10), started with a training session on the traditional box trainer for half an hour followed by a session on the SimSurgery VR simulator for half an hour; group B (n = 10), started with the same session on the SimSurgery VR simulator, followed by the session on the traditional box trainer. Finally, suturing and knot-tying skills were assessed by an expert observer, using a standard evaluation form. The same was done after the initial training on the box in group A, as a control. Overall, the total score of group A was higher than both group B and control. All the participants scored the features of the box trainer significantly higher than those of the VR simulator (p < 0.001), 46.7% was of the opinion that the box alone would be sufficient for laparoscopic suturing training. The authors concluded that VR simulation does not have a significant additional value in laparoscopic suturing training, over traditional box trainers.

Cohen et al. (2006) conducted an RCT (n = 45) to determine whether a 10-hour structured training program that used the GI Mentor simulator provided an objective benefit to novice gastroenterology fellows before performing real colonoscopies. Subjects were randomized to receive 10 hours of unsupervised training on the GI Mentor or no simulator experience during the first 8 weeks of fellowship. After this period, both groups began performing real colonoscopies. Study measured the mixed-effects model comparison between the 2 groups of objective and subjective competency scores and patient discomfort in the performance of real colonoscopies over time.
Results from the study showed that Fellows in the simulator group had significantly higher objective competency rates during the first 100 cases. A mixed-effects model demonstrated a higher objective competence overall in the simulator group (P <.0001), with the difference between groups being significantly greater during the first 80 cases performed. The median number of cases needed to reach 90% competency was 160 in both groups. The patient comfort level was similar. The authors concluded that fellows who underwent GI Mentor training performed significantly better during the early phase of real colonoscopy training.

Eversbusch et al. (2004) analyzed the learning curve for the GI Mentor II endoscope trainer and assessed the contribution of psychomotor training for an improvement in the performance of virtual colonoscopy. 28 subjects were divided into three groups on the basis of their experience with gastrointestinal (GI) endoscopy: experienced surgeons (group 1, n = 8)) residents (group 2, n = 10); and medical students (group 3, n = 10)). The participants were tested on the GI Mentor II virtual reality simulator 10 consecutive times. Assessment of the learning curve was based on time used, number of punctured balloons, and number of wall collisions. In the second part of the study, 20 subjects who had never performed GI endoscopy were included. After performing a virtual colonoscopy, they were randomized to a group that received psychomotor training and a control group. Finally, all subjects performed a virtual colonoscopy. The study found that the learning curve for time spent reached a plateau after the second repetition for group 1 (p < 0.05), after the fifth repetition for group 2 (p < 0.05), and after the seventh repetition for group 3 (p < 0.05). Experienced surgeons did not improve their scores for regarding number of balloons punctured or number of wall collisions (p > 0.05), indicating the absence of a learning curve for these parameters. Group 2 improved their scores up to the fourth and fifth repetitions, respectively (p < 0.05), and group 3 up to the fifth and seventh repetitions, respectively (p < 0.05).
Experienced surgeons achieved the best performance, followed by group 2 and then group 3. The surgeons who had received psychomotor training performed the second virtual colonoscopy significantly faster than the control group (p < 0.001) and made significantly greater improvement in all other parameters as well. Investigators concluded that there were different learning curves for surgeons depending on their endoscopic background and psychomotor training had a significant effect on the performance of a simulated colonoscopy.

Grantcharov et al. (2004) conducted an RCT among surgical trainees (n = 20) to test the impact of virtual reality (VR) surgical simulation on improvement of psychomotor skills relevant to the laparoscopic cholecystectomy. After performing laparoscopic cholecystectomy on patients in the operating room (OR), the participants were randomized to receive VR training (ten repetitions of all six tasks on the Minimally Invasive Surgical Trainer – Virtual Reality (MIST-VR)) or no training. Subsequently, all subjects performed a further laparoscopic cholecystectomy in the OR. Time to complete the procedure, error score and economy of movement score were assessed during the laparoscopic procedure in the OR. The results indicated that no differences in baseline variables between the two groups. Surgeons who received VR training performed laparoscopic cholecystectomy significantly faster than the control group (P = 0.021). Furthermore, those who had VR training showed significantly greater improvement in error (P = 0.003) and economy of movement (P = 0.003) scores. Based on these observations, the authors concluded that surgeons who received VR simulator training showed significantly greater improvement in performance in the OR than those in the control group and therefore, a valid tool for training of laparoscopic psychomotor skills. Thus, VR simulator training could be incorporated into surgical training programmes.
Larsen et al. (2009) conducted a prospective randomised controlled and blinded trial to assess the effect of virtual reality training on an actual laparoscopic operation. This study involved first and second year registrars specialising in gynaecology and obstetrics (n = 24). The intervention group undertook a specific training programme in the simulator (LapSim Gyn; Surgical Science, Gothenburg, Sweden) and the control group continued standard clinical education. Results from the study showed that simulator trained group (n=11) reached a median total score of 33 points, equivalent to the experience gained after 20-50 laparoscopic procedures, whereas the control group (n=10) reached a median total score of 23 points, equivalent to the experience gained from fewer than five procedures (P<0.001). The median total operation time in the simulator trained group was 12 minutes and in the control group was 24 minutes (P<0.001). The authors commented that skills in laparoscopic surgery can be increased in a clinically relevant manner using proficiency based virtual reality simulator training and therefore, simulator training should be considered before trainees carry out laparoscopic procedures.

Neequaye et al. (2007) conducted an RCT to determine the nature of skills acquisition on the renal and iliac modules of a commercially-available VR simulator. 20 surgical trainees without endovascular experience were randomised to complete eight sessions on a VR iliac (group A) or renal (group B) training module. To determine skills transferability across the two procedures, all subjects performed two further VR cases of the other procedure. Performance was recorded by the simulator for parameters such as time taken, contrast fluid usage and stent placement accuracy. They found that, during training, both groups demonstrated statistically significant VR learning curves: group A for procedure time (p < 0.001) and stent placement accuracy (p = 0.013) group B for procedure time (p < 0.001), fluoroscopy time (p = 0.003) and volume of contrast fluid used (p < 0.001).
At crossover, subjects in group B (renal trained) performed to the same level of skill on the simulated iliac task as group A. However, those in group A (iliac trained) had a significantly higher fluoroscopy time (median 118 vs. 72 seconds, \( p = 0.020 \)) when performing their first simulated renal task than for group B. Overall, the results from study indicate that novice endovascular surgeons can significantly improve their performance of simulated procedures through repeated practice on VR simulators. Although, study demonstrated transference of skills between VR endovascular skills tasks, for complex task training, such as selective arterial cannulation in simulators and possibly in the real world appears to involve a separate skill.

**Rowe & Cohen (2002)** tested the effectiveness of AccuTouch Flexible Bronchoscopy Simulator to teach clinicians the psychomotor skills necessary to use the fiberoptic bronchoscope. Paediatric residents with no prior experience in fiberoptic bronchoscopy were studied. All residents performed two fiberoptic intubations on patients during the study. The Simulator group (\( n = 12 \)) received training on the Simulator between their cases, whereas the Control group (\( n = 8 \)) did not. The results from the study showed that significant improvement in time to completion of endotracheal intubation, as well as other performance indicators. Time to completion of successful intubation with a bronchoscope was reduced from 5.15 to 0.88 min (\( P < 0.001 \)). The number of times that the tip of the bronchoscope hit the mucosa was reduced from 21.4 to 3.0 (\( P < 0.001 \)). The amount of time that the resident spent viewing the mucosa decreased from 2.24 to 0.19 min (\( P < 0.001 \)). The percent of time viewing the channel of the airway increased from 58.5% to 80.4% (\( P = 0.004 \)). The authors concluded that the simulator was successful at training residents the psychomotor skills necessary for fiberoptic intubation.
Shirai et al. (2008) conducted an RCT on the effectiveness of a computer-based simulator for basic training in esophagogastroduodenoscopy (EGD?). The simulator group (n = 10) received 5 h of training with the GI-Mentor II plus bedside training, while the non-simulator group (n = 10) received bedside training. Subsequently, each subject performed endoscopy twice for assessment. The scores were significantly higher in the simulator group with respect to insertion of the endoscope into the oesophagus, passing the EGJ into the antrum, passing through the pyloric ring, and examination of the duodenal bulb and the fornix. There was no significant difference in the total procedure time between the simulator group and the non-simulator group (14:40 min vs. 14:05 min). Overall, results from the study indicate that the GI-Mentor II was more effective with regard to the items related to manipulation skills.

Verdaasdonk et al. (2008) conducted an RCT on the efficacy of transfer validity of knot-tying training on a virtual-reality (VR) simulator to a realistic laparoscopic environment. Group A (the experimental group), received additional training with the knot-tying module on the simulator and group B (controls) did not receive additional manual training. The results indicated that trainees in group A (n = 9) were significantly faster than the controls (n = 10), with a median of 262 versus 374 seconds (p = 0.034). Group A made a significantly lower number of errors than the controls (median of 24 versus 36 errors, p = 0.030). Based on this, authors concluded that the VR simulator under study can provide effective training of knot-tying skills.

Summary of other types of reviews

Seymour (2008) examined the background, results, and significance of randomized trials that have been undertaken to demonstrate that skills acquired during VR training transfer to the operating room. Of the seven published studies of laparoscopic skills transfer that were
reviewed, one failed to demonstrate transfer of skills. Several skills transfer studies of VR flexible endoscopic trainers have been reviewed and the results were similar to those reported for laparoscopic VR training. Measures of performance, time to procedure completion, and achievement of specific procedural goals have been used as clinical skill metrics. However, some of these studies have also examined clinical outcomes, including patient discomfort and satisfaction, as metrics for effectiveness of training.

Cakmak et al. (2005) introduced VSOne as a multi-media based medical e-learning system combined with a Virtual Reality based haptic training system for laparoscopic surgery. This paper gave a system overview with a detailed description of the hardware and software. Special emphasis has being given to modelling with the authoring software KisMo, which enables to create patient-specific simulation models. Results were presented for simple dexterity training modules and full surgical procedure simulation for cholecystectomy and gynaecology. The paper closed with an evaluation of VSOne in a comparative study with classic training methods and with a discussion on the benefits of VR based training systems.

Results from quasi-experimental studies

Aggarwal and Black (2006) used two group comparison study to assess the effectiveness of virtual reality (VR) simulation for endovascular training of surgeons inexperienced in this technique. Twenty consultant vascular surgeons were divided into those who had performed > 50 endovascular procedures as primary operator (n = 8), and those having performed < 10 procedures (n = 12). To test for endovascular skill rather than procedural knowledge, all subjects performed a renal artery balloon angioplasty and stent procedure. Surgeons with endovascular skills performed two repetitions and those without completed six repetitions of the same task. The simulator recorded time taken for the procedure, the amount of contrast
fluid used and total fluoroscopy time. Finding from the study showed that over the six sessions, the inexperienced group made significant improvements in performance for time taken (p = 0.007) and contrast fluid usage (p = 0.021), achieving similar scores at the end of the training program to the experienced group. The authors concluded that VR simulation may be useful for the early part of the learning curve for surgeons who wish to expand their endovascular interests.

Aggarwal and Tully (2006) conducted a prospective cohort study to evaluate a VR ectopic module (LAPSIM; Surgical Science, Gothenburg, Sweden), in terms of its validity as a training and assessment tool for gynaecological surgeons, and specifically whether it could viably be integrated into a skills curriculum. Thirty gynaecological surgeons were divided into novice (<10 laparoscopic procedures), intermediate (20–50) and experienced (>100) groups. All subjects performed ten repetitions of the virtual ectopic pregnancy module, in a distributed manner. Operative performance was assessed by the time taken to perform surgery, blood loss and total instrument path length. Findings from the study showed that significant differences between the groups at the second repetition of the ectopic module for time taken (median 551.1 versus 401.2 versus 249.2 seconds, P = 0.001), total blood loss (median 304.2 versus 187.4 versus 123.3 ml, P = 0.031) and total instrument path length (median 17.8 versus 8.3 versus 6.8 m, P = 0.023). Overall, results from the study showed that Gynaecological surgeons with minimal laparoscopic experience can improve their skills during short-phase training on a VR procedural module. In contrast, experienced operators showed non-significant improvements. The authors concluded that VR simulation may be useful for the early part of the learning curve for surgeons who wish to learn to perform laparoscopic salpingectomy for ectopic pregnancy.
**Gallagher et al. (2004)** conducted a two group comparison study on 100 laparoscopic novices to study the effectiveness of Minimally Invasive Surgical Trainer in Virtual Reality (MIST-VR) as an assessment tool using criteria levels based on expert performance. The performance of 100 laparoscopic novices was compared to that of 12 experienced and 12 inexperienced laparoscopic surgeons. Each subject completed six tasks on the MIST-VR three times. The outcome measures were time to complete the task, number of errors, economy of instrument movement, and economy of diathermy. After three trials, investigators reported that the mean performance of the medical students approached that of the experienced surgeons. However, 7–27% of the scores of the students fell more than two SD below the mean scores of the experienced surgeons (the criterion level). The investigators concluded that MIST-VR system was capable of evaluating the psychomotor skills necessary in laparoscopic surgery and discriminating between experts and novices.

**Knoll et al. (2005)** conducted a two group comparison study to evaluate UroMentor (Simbionix Ltd, Israel) trainer. Twenty experienced urologists were monitored during simulated flexible URS for treating a lower calyceal stone. A further five urological residents with no endourological experience were trained on the UroMentor in rigid URS for ureteric stone treatment. Their acquired clinical skills were subsequently compared to those of five urological residents who received no simulator training. The results achieved by urologists highly experienced in flexible URS were clearly better than those of the novices. (Operation time $P<0.05$). Conversely, standardized simulator training enabled inexperienced residents to become competent in semi-rigid URS. ($P<0.05$) Overall, the results of the study supported the value of simulator training for URS on clinical performance.

**Langelotz et al. (2005)** conducted a two group comparison study to validate LapSim laparoscopic simulator as an assessment and educational tool in surgery.
Participants (n = 115) were stratified into two groups based on their laparoscopic experience. All subjects completed a laparoscopic training module consisting of five different exercises. The time to perform each task was measured, as were the path lengths of the instruments and their economy of the movements. Surgeons with less experience needed more time for completion of the exercises (P < 0.01) and instrument movements were less economic (P<0.01) than group with more experience. They also showed longer path lengths (each instrument P<0.05). The authors concluded that laparoscopic simulator can serve as an instrument for the assessment of experience in laparoscopic surgery. This study has added the aspect that virtual laparoscopy furthermore offers to reliably assess basic laparoscopic abilities in inexperienced as well as experienced surgeons.

**Schijven et al. (2005)** conducted a two group comparison study to investigate the operation performance of surgical residents after their participating in the virtual reality laparoscopic cholecystectomy training course. Course participants’ (n = 12) operating room performance was compared with the operating room performance of a matched control group (n = 12). The results from study suggested that Virtual Reality laparoscopic cholecystectomy training course improves surgical skill in the operating room above the level of residents trained by a variety of other training methods (p-value 0.004 and 0.013).

**Gomoll et al. (2008)** conducted a 3-year follow-up study to assess the effectiveness of surgical simulation as an important tool for the evaluation of surgical skills. This study investigated the performances of 10 orthopaedic residents who were retested 3 years after initial evaluation on a VR simulator for shoulder arthroscopy. The results of the study showed that subjects improved significantly (P < .02 for all) in the 4 simulator parameters: completion time (~51%), probe collisions (~29%), average velocity (+122%), and distance
travelled (−32%). The investigators concluded that these results are further validating the use of VR simulation for the evaluation of surgical skills.

Observational studies

Goldmann et al. (2006) conducted an observational study (n = 11) to test the effectiveness of virtual reality (VR) airway simulator (the AccuTouch Virtual Reality Bronchoscopy Simulator) to teach residents basic fiberoptic intubation (FOI) skills. Time to intubation before and after a 4-day training period using an adult VR FOI scenario and time to intubation using a fresh human cadaver two weeks after the training experience were measured. The results of the study showed that residents were able to significantly improve time to intubation in the VR scenario (114 vs 75 seconds; P = 0.001). Novices differed from experienced attending anaesthesiologists in time to intubation in the VR scenario, before but not after training (114 vs. 79 seconds compared with 75 vs. 72 seconds). Furthermore, there was no difference in time to intubation in the cadaver between trained novices and experienced attending anaesthesiologists (24 vs. 23 seconds; P > 0.05). The authors concluded that use of a VR airway simulator enables anaesthesia residents to acquire basic FOI skills comparable to those of experienced anaesthesiologists in a human cadaver.

Studies with pre and post test evaluations

Hassan et al. (2006) conducted a pre-post intervention study to determine if a three-day laparoscopic skills course can improve laparoscopic skills of residents measured by a virtual reality laparoscopy-simulator (LapSim). 44 participants with various degree of experience in laparoscopic surgery and 6 consultants attending as tutors of the course (gold standard) were recruited as subjects. 20 medical students in their final year were chosen as a second control group. LapSim was used to assess laparoscopic skills of participants before and after the
course. The results of the study showed that advanced participants of the test group completed the task significantly faster (p = 0.019), with smaller error score (p = 0.023), and more economy of motion [path length (p = 0.014) and angular path (p = 0.049)] than before the course. The novices of the test group and both control groups showed no significant improvement of their performance parameters (p >0.05). Investigators concluded that residents with some degree of experience in laparoscopic surgery profit mostly from laparoscopic skills courses.

Grantcharov et al. (2001) conducted a pre-post intervention study on 14 surgical residents to validate the role of virtual reality computer simulation as a method for evaluating surgical laparoscopic skills. On day 1, they performed two runs of all six tasks on the Minimally Invasive Surgical Trainer, Virtual Reality (MIST-VR). On day two, they performed a laparoscopic cholecystectomy on living pigs; afterward, they were tested again on the MISTVR. The results of this study suggest that laparoscopic performance in the animal model correlated significantly with performance on the computer simulator. There were significant correlations between error scores in vivo and three of the six in vitro tasks (p < 0.05). In vivo economy scores correlated significantly with noneconomic right-hand scores for five of the six tasks and with non-economy left-hand scores for one of the six tasks (p < 0.05). They concluded that the computer model seems to be a promising objective method for the assessment of laparoscopic psychomotor skills.

Torkington et al. (2001) conducted a pre-post intervention study on 13 Basic Surgical Trainees to assess the Basic Surgical Skills course as a vehicle for teaching basic laparoscopic surgical skills to BST and to ascertain the length of time that new skills were retained. The virtual reality simulator MIST-VR was used to assess 13 trainees before and after the course and again 3 weeks and 3 months later.
A control group of senior medical students (n =13) with no laparoscopic experience and no formal training were assessed at the time points equivalent to the subject group. Investigators found that all parameters improved significantly after the course, with the exception of distance travelled by the instruments. All outcome measures were significantly improved at 3 weeks. The control group showed a nonsignificant trend toward improvement in all parameters. The results of this study suggest that the Basic Surgical Skills course produces quantifiable improvements in laparoscopic skill that are measurable by MIST-VR.
<table>
<thead>
<tr>
<th>Study design</th>
<th>Author (year)</th>
<th>Application</th>
<th>Population</th>
<th>Outcome studied</th>
<th>Effect estimate or difference obtained</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meta-analysis</td>
<td>Gurusamy et al (2008)</td>
<td>VR stimulator and video trainer (VT)</td>
<td>Surgical trainees, medical students and surgical residents</td>
<td>Time taken to perform the evaluation task on the simulation model (after training), operating time error score, the number of undesirable movements, improvement in task performance and participant satisfaction.</td>
<td>This review included a methodical search of all the relevant literature databases and included randomized trials that evaluated the effectiveness of VR training. The results showed in trainees without surgical experience VR training decreased the time taken to complete a task (SMD = 1.09 (95% CI = 1.50 to −0.68), increased accuracy and decreased errors compared with no training.</td>
<td>There is convincing evidence that VR training is a useful supplement to standard laparoscopic training in surgical residents with limited laparoscopic experience.</td>
</tr>
<tr>
<td>Systematic reviews</td>
<td>Aggarwal et al (2004)</td>
<td>Box trainer with either innate models or animal tissues. MIST-VR laparoscopic simulator (Mentice, Gothenburg, Sweden) LapSim (Surgical Science, Gothenburg, Sweden).</td>
<td>Experienced, inexperienced and novice laparoscopic surgeons</td>
<td>Performance task on the simulator (time spend, error and economy of movement scores) Analysis of learning curves of experts and novices</td>
<td>Box trainers lack objective assessment of skill acquisition. Virtual reality simulators have the ability to teach laparoscopic psychomotor skills, and objective assessment is now possible using dexterity-based and video analysis systems.</td>
<td>There is sufficient evidence for incorporation of virtual simulation into current training programmes.</td>
</tr>
</tbody>
</table>
Table 2: Included studies: brief characteristics and main findings (continued)

<table>
<thead>
<tr>
<th>Study design</th>
<th>Author (year)</th>
<th>Application</th>
<th>Population</th>
<th>Outcome studied</th>
<th>Effect estimate or difference obtained</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized Controlled Trials</td>
<td>Botden et al. (2008)</td>
<td>SimSurgery VR simulator systems (Oslo, Norway) and traditional box trainers</td>
<td>Surgical and gynaecology residents</td>
<td>Suturing and knot-tying skills were assessed</td>
<td>No significant differences between the ratings of the assessment after laparoscopic suturing training on only a traditional box trainer (control) or the combination with a VR simulator (groups A and B); (means of 30.80, 27.60, 28.20)</td>
<td>Study showed that VR simulation does not have a significant additional value in laparoscopic suturing training, over traditional box trainers.</td>
</tr>
<tr>
<td></td>
<td>Cohen et al. (2006)</td>
<td>GI Mentor</td>
<td>First-year GI fellows randomized into simulator training and no training.</td>
<td>Comparison of objective and subjective competency scores and patient discomfort in the performance of real colonoscopies</td>
<td>Fellows who underwent GI Mentor training performed significantly better during the early phase of real colonoscopy training. (Higher objective competence overall in the simulator group ($P &lt; .0001$).)</td>
<td>Study results showed a benefit of training on the GI Mentor before performing real colonoscopies.</td>
</tr>
</tbody>
</table>
Table 2: Included studies: brief characteristics and main findings (continued)

<table>
<thead>
<tr>
<th>Study design</th>
<th>Author (year)</th>
<th>Application</th>
<th>Population</th>
<th>Outcome studied</th>
<th>Effect estimate or difference obtained</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eversbusch et al. (2004)</td>
<td>GI Mentor II virtual reality simulator</td>
<td>Experienced surgeons (GI endoscopy), residents and medical students</td>
<td>Assessment of the learning curve. Based on: time used, number of punctured balloons, and number of wall collisions.</td>
<td>Results showed significant differences in the familiarization curves on the simulator among subjects of three experience levels.</td>
<td>There were different learning curves for surgeons depending on their endoscopic background. Psychomotor training had a significant effect on the performance of a simulated colonoscopy.</td>
</tr>
<tr>
<td></td>
<td>Grantcharov et al. (2004)</td>
<td>Minimally Invasive Surgical Trainer – Virtual Reality (MIST-VR)</td>
<td>Surgical trainees randomized to receive VR training and no training.</td>
<td>Time to complete the procedure, error score and economy of movement score were assessed during the laparoscopic procedure in the operation room.</td>
<td>Surgeons who received VR simulator training showed significantly greater improvement in time ((P = 0.021)), error ((P = 0.003)) and economy of movement ((P = 0.003)) scores in the OR than those in the control group.</td>
<td>VR surgical simulation can be use as a valid tool for training of laparoscopic psychomotor skills.</td>
</tr>
</tbody>
</table>
Table 2: Included studies: brief characteristics and main findings (continued)

<table>
<thead>
<tr>
<th>Study design</th>
<th>Author (year)</th>
<th>Application</th>
<th>Population</th>
<th>Outcome studied</th>
<th>Effect estimate or difference obtained</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Larsen et al. (2009)</td>
<td>LapSim Gyn v 3.0.1; Surgical Science, Gothenburg, Sweden</td>
<td>First and second year registrars specialising in gynaecology and obstetrics.</td>
<td>Technical performance and operation time</td>
<td>The simulator trained group reached a median total score of 33 points whereas the control group reached a median total score of 23 (P&lt;0.001). Median total operation time was 12 min. compared to 24 min. (P&lt;0.001).</td>
<td>Skills in laparoscopic surgery can be increased in a clinically relevant manner using proficiency based virtual reality simulator training. Thus, simulator training can be incorporated into surgical training programmes.</td>
</tr>
<tr>
<td></td>
<td>Neequaye et al. (2007)</td>
<td>Vascular Intervention System Training simulator (VIST, Mentice Corporation, Goteborg, Sweden)</td>
<td>Surgical trainees without endovascular experience</td>
<td>Performance parameters (time taken, contrast fluid usage and stent placement accuracy).</td>
<td>During training, both (VR iliac or renal training module) groups demonstrated statistically significant VR learning curves. At crossover, subjects in renal trained performed to the same level of skill on the simulated iliac task.</td>
<td>Novice endovascular surgeons can significantly improve their performance of simulated procedures through repeated practice on VR simulators.</td>
</tr>
</tbody>
</table>
Table 2: Included studies: brief characteristics and main findings (continued)

<table>
<thead>
<tr>
<th>Study design</th>
<th>Author (year)</th>
<th>Application</th>
<th>Population</th>
<th>Outcome studied</th>
<th>Effect estimate or difference obtained</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rowe &amp; Cohen</td>
<td>AccuTouch</td>
<td>Paediatric</td>
<td>Performance of</td>
<td>Significant improvement was seen in time to completion of endotracheal intubation ($P&lt;0.001$) as well as other performance indicators.</td>
<td>Bronchoscopy simulator was very effective in teaching residents the psychomotor skills necessary for fiberoptic intubation.</td>
</tr>
<tr>
<td>(2002)</td>
<td>flexible</td>
<td>bronchoscope simulator</td>
<td>residents without prior experience in bronchoscopy.</td>
<td>fiberoptic intubation were analyzed for: time to visualization of the carina, and number and time that the bronchoscope tip hit the mucosa</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shirai et al.</td>
<td>GI-Mentor II</td>
<td>Medical</td>
<td>Endoscopy</td>
<td>The performance of endoscopy was improved by simulator training. The scores were significantly higher in the simulator group regarding insertion of the endoscope into the oesophagus. There was no significant difference in the total procedure time.</td>
<td>The simulator was more effective with regard to the items related to manipulation skills. Computer-based simulator training in EGD is useful for beginners.</td>
</tr>
<tr>
<td>(2008)</td>
<td>simulator</td>
<td>residents randomized into a simulator group and a non-simulator group</td>
<td>performance was evaluated according to a five-grade scale for a total of 11 items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td>Author (year)</td>
<td>Application</td>
<td>Population</td>
<td>Outcome studied</td>
<td>Effect estimate or difference obtained</td>
<td>Conclusion</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------------</td>
<td>------------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td></td>
<td>Verdaasdonk et al. (2008)</td>
<td>The SIMENDO endoscopic simulator (Delltatech, Delft, the Netherlands)</td>
<td>First- and second-year surgical trainees randomly allocated into experimental and control groups</td>
<td>Assessment of video recordings: Time taken to tie the knot and number of predefined errors. Subjective assessments were also made using a global rating list.</td>
<td>Surgical trainees who received knot-tying training on the VR simulator were faster with a median of 262 versus 374 seconds (p = 0.034) and made fewer errors (median of 24 versus 36 errors, p = 0.030) than the controls.</td>
<td>The VR simulator under study is a useful tool to train laparoscopic knot-tying. VR simulator training may maximize the efficiency of instruction from experts.</td>
</tr>
<tr>
<td>Narrative reviews</td>
<td>Seymour (2008)</td>
<td>MIST-VR LapSim</td>
<td>Medical students and surgical residents</td>
<td>Operating room performance</td>
<td>This review examined the background, results, and significance of randomized trials that have been undertaken to demonstrate that skills acquired during VR training transfer to the operating room.</td>
<td>Of the seven published studies of laparoscopic skills transfer that were reviewed, one failed to demonstrate transfer of skills. Several skills transfer studies of VR flexible endoscopic trainers have been reviewed and the results were similar to those reported for laparoscopic VR training.</td>
</tr>
<tr>
<td>Study design</td>
<td>Author (year)</td>
<td>Application</td>
<td>Population</td>
<td>Outcome studied</td>
<td>Effect estimate or difference obtained</td>
<td>Conclusion</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------------</td>
<td>---------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td></td>
<td>Cakmak et al (2005)</td>
<td>VSOne - Virtual Endoscopic Surgery Training (VEST) system</td>
<td>Medical students and laparoscopic surgeons</td>
<td>Testing the skill transferability</td>
<td>This paper introduced VSOne as a fully-featured VR training system for laparoscopic surgery with a system overview of VSOne. Detailed description of the hardware and software given. Evaluation of VSOne done in comparative study with classic training methods. Discussion on the benefits of VR based training systems given.</td>
<td>Surgical training centres have recognized the importance and the potential of VR based training systems. VR based training will become an integral part of surgical education and training programmes.</td>
</tr>
<tr>
<td>Quasi-experimental studies</td>
<td>Aggarwal and Black (2006)</td>
<td>Vascular Interventional Surgical Trainer (VIST simulator)</td>
<td>Surgeons with extensive experience in open vascular surgical procedures divided into two groups, based upon their experience in endovascular procedures.</td>
<td>Performance evaluation (total time taken, total amount of contrast fluid used and fluoroscopy time)</td>
<td>Surgeons with minimal endovascular experience can improve their time taken (p=0.007) and contrast fluid usage (p=0.021), during short-phase training on a VR endovascular task.</td>
<td>VR simulation may be useful for the early part of the learning curve for surgeons who wish to expand their endovascular interests.</td>
</tr>
</tbody>
</table>
Table 2: Included studies: brief characteristics and main findings (continued)

<table>
<thead>
<tr>
<th>Study design</th>
<th>Author (year)</th>
<th>Application</th>
<th>Population</th>
<th>Outcome studied</th>
<th>Effect estimate or difference obtained</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aggarwal and Tully (2006)</td>
<td>VR ectopic module (LAPSIM; Surgical Science, Gothenburg, Sweden).</td>
<td>Gynaecological Surgeons divided into 3 groups according to their experience</td>
<td>Operative performance was assessed by the time taken to perform surgery, blood loss and total instrument path length.</td>
<td>Experts performed significantly better than the intermediates and in turn the novices. Novices managed to achieve similar levels of skill towards the end of the training period.</td>
<td>VR simulation may be useful for the early part of the learning curve for surgeons who wish to learn to perform laparoscopic salpingectomy for ectopic pregnancy.</td>
</tr>
<tr>
<td></td>
<td>Gomoll et al. (2008)</td>
<td>Procedicus virtual reality arthroscopy simulator</td>
<td>Orthopaedic residents</td>
<td>Time to completion, number of probe collisions with the tissues, average probe velocity, and distance travelled with the tip of the simulated probe compared to an optimal computer-determined distance.</td>
<td>Subjects significantly improved ($P &lt; .02$ for all) their performance on simulator retesting 3 years after initial evaluation. Groups with similar arthroscopic experience consistently demonstrate equivalent scores on the simulator.</td>
<td>Results further validate the use of surgical simulation as an important tool for the evaluation of surgical skills.</td>
</tr>
<tr>
<td>Study design</td>
<td>Author (year)</td>
<td>Application</td>
<td>Population</td>
<td>Outcome studied</td>
<td>Effect estimate or difference obtained</td>
<td>Conclusion</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>-------------</td>
<td>------------</td>
<td>----------------</td>
<td>-----------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td></td>
<td>Gallagher et al. (2004)</td>
<td>Minimally Invasive Surgical Trainer in Virtual Reality (MIST-VR)</td>
<td>Laparoscopic surgeons with different levels of experience</td>
<td>Performance measures: time to complete the task, number of errors, economy of instrument movement, and economy of diathermy</td>
<td>After three trials, the mean performance of the medical students approached that of the experienced surgeons. However, 7–27% of the scores of the students fell more than two SD below the mean scores of the experienced surgeons (the criterion level).</td>
<td>The MIST-VR system is capable of evaluating the psychomotor skills necessary in laparoscopic surgery and discriminating between experts and novices.</td>
</tr>
<tr>
<td></td>
<td>Knoll et al. (2005)</td>
<td>UroMentor (Simbionix Ltd, Israel)</td>
<td>Urologists (with and without endourological experience)</td>
<td>Performance score (total operation time, stone contact time, and complications)</td>
<td>The results achieved by urologists highly experienced in flexible URS were clearly better than those of the novices. (Operation time $P&lt;0.05$). Conversely, standardized simulator training enabled inexperienced residents to become competent in semi-rigid URS. ($P&lt;0.05$)</td>
<td>Overall, the results of the study supported the value of simulator training for URS on clinical performance.</td>
</tr>
</tbody>
</table>
Table 2: Included studies: brief characteristics and main findings (continued)

<table>
<thead>
<tr>
<th>Study design</th>
<th>Author (year)</th>
<th>Application</th>
<th>Population</th>
<th>Outcome studied</th>
<th>Effect estimate or difference obtained</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Langelotz et al. (2005)</td>
<td>LapSim</td>
<td>Surgeons with and without laparoscopic experience</td>
<td>Laparoscopic skills (time to perform, economy of the movements and the path lengths)</td>
<td>Surgeons with less experience needed more time for completion of the exercises (P &lt; 0.01) and instrument movements were less economic (P&lt;0.01) than group with more experience. They also showed longer path lengths (each instrument P&lt;0.05).</td>
<td>Laparoscopic simulator can serve as an instrument for the assessment of experience in laparoscopic surgery.</td>
</tr>
<tr>
<td></td>
<td>Schijven et al. (2005)</td>
<td>Xitact LS500 laparoscopy simulator platform (Xitact SA, Morges, Switzerland)</td>
<td>Novice urological residents</td>
<td>Operating room performance (fluency, judgement and carefulness)</td>
<td>VR training group perform significantly better. (p - value 0.004 and 0.013). They valued their course highly in terms of their laparoscopic surgical skills.</td>
<td>Virtual Reality laparoscopic cholecystectomy training course improves surgical skill in the operating room above the level of residents trained by a variety of other training Methods.</td>
</tr>
</tbody>
</table>
Table 2: Included studies: brief characteristics and main findings (continued)

<table>
<thead>
<tr>
<th>Study design</th>
<th>Author (year)</th>
<th>Application</th>
<th>Population</th>
<th>Outcome studied</th>
<th>Effect estimate or difference obtained</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observational studies</td>
<td>Goldmann et al. (2006)</td>
<td>AccuTouch Virtual Reality Bronchoscopy Simulator</td>
<td>Experienced anaesthesiologists, and residents who did not have any clinical experience in fiberoptic intubation (FOI)before</td>
<td>Time to intubation before and after a training period using an adult VR FOI scenario and time to intubation using a fresh human cadaver two weeks after the training experience</td>
<td>Novices who had been trained with the simulator performed significantly faster in the cadaver than novices who had not (24 vs. 86 seconds; $P &lt; 0.001$). There was no difference in time to intubation in the cadaver between trained novices and experienced anaesthesiologists (24 vs. 23 seconds; $P &gt; 0.05$).</td>
<td>Use of a VR airway simulator enables anaesthesia residents to acquire basic FOI skills comparable to those of experienced anaesthesiologists in a human cadaver.</td>
</tr>
</tbody>
</table>
Table 2: Included studies: brief characteristics and main findings (continued)

<table>
<thead>
<tr>
<th>Study design</th>
<th>Author (year)</th>
<th>Application</th>
<th>Population</th>
<th>Outcome studied</th>
<th>Effect estimate or difference obtained</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before-after intervention studies</td>
<td>Hassan et al. (2006)</td>
<td>LapSim, Surgical Science Ltd., Goteborg/ Sweden</td>
<td>Surgical residents (Advanced training and novices)</td>
<td>Time to complete the tasks, error score, and economy of motion parameters.</td>
<td>Advanced participants of the test group completed the task significantly faster (p = 0.019), with smaller error score (p = 0.023), and more economy of motion [path length (p = 0.014) and angular path (p = 0.049)] than before the course. The novices showed no significant improvement of their performance (p &gt; 0.05)</td>
<td>Residents with some degree of experience in laparoscopic surgery excluding novices profit mostly from laparoscopic skills courses by a virtual reality simulator.</td>
</tr>
<tr>
<td></td>
<td>Grantcharov et al. (2001)</td>
<td>Minimally Invasive Surgical Trainer, Virtual Reality (MIST-VR)</td>
<td>Surgical residents</td>
<td>Trainees’ performance on the animal operation, giving scores for total performance error and economy of motion.</td>
<td>Scores for tasks performed on a computer simulation system (MISTVR) are comparable to performance during operations on living animals.</td>
<td>Computer model seems to be a promising objective method for the assessment and evaluation of laparoscopic psychomotor skills.</td>
</tr>
</tbody>
</table>
Table 2: Included studies: brief characteristics and main findings (continued)

<table>
<thead>
<tr>
<th>Study design</th>
<th>Author (year)</th>
<th>Application</th>
<th>Population</th>
<th>Outcome studied</th>
<th>Effect estimate or difference obtained</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Torkington et al. (2001)</td>
<td>Virtual reality simulator MIST-VR (Virtual Presence, London, SE1 2NL, UK)</td>
<td>Basic Surgical Trainees (VR training)</td>
<td>Measures of distance travelled, distance efficiency ratio, time taken, number of errors made, and number of movements made in completing a virtual laparoscopic task.</td>
<td>All outcome measures were significantly improved at 3 weeks. Time (p &lt; 0.01), The distance economy (p &lt; 0.05), Errors (p &lt; 0.05), Movements made (p &lt; 0.01). The control group showed a non-significant trend toward improvement in all parameters.</td>
<td>The Basic Surgical Skills course produces quantifiable improvements in laparoscopic skill that are measurable by MIST-VR.</td>
</tr>
<tr>
<td>Type of intervention</td>
<td>Application title (Platforms)</td>
<td>Author</td>
<td>Studies that have used it</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>------------------------------</td>
<td>--------</td>
<td>--------------------------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training of health professionals</td>
<td>Vascular Interventional Surgical Trainer (VIST simulator)</td>
<td>Mentice Corporation, Gothenburg, Sweden</td>
<td>Aggarwal and Black (2006) Neequaye et al. (2007)</td>
<td>VIST System is a high fidelity endovascular simulator. Unique tactile response enabled by active force feedback with independent translation and rotation (patented technology). The training system consists of the software simulation of the cardiovascular system, the haptic interface device, the instructional system and two monitors, one for the synthetic x-ray and one for the instructional system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training of health professionals</td>
<td>Surgical Education Platform (SEP) trainer</td>
<td>Software: SEP 1.04.3 SimSurgery, Oslo, Norway Hardware platform: 1. SimSurgery, Oslo, Norway 2. Xitact/Mentice SA, Morges, Switzerland</td>
<td>Botden et al. (2008)</td>
<td>Two SimSurgery VR simulator systems were used in this study. Both VR simulator systems ran the SimPort software. However, their hardware platform differed: one VR simulator system incorporated the SimPack platform, while the other system made use of two Xitact HTP instrument ports. The SimPort software package used in this study contains various training modules, including exercises related to specific laparoscopic skills, techniques, or procedures. However, software does not support haptic features.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment of skills</td>
<td>VSOne (Virtual Endoscopic Surgery Training (VEST) system)</td>
<td>Institute for Applied Informatics, IAI at the Forschungszentrum Karlsruhe in close cooperation with Select IT-VEST Systems AG (Bremen, Germany)</td>
<td>Cakmak et al (2005)</td>
<td>The simulation system is based on a multiprocessor graphics workstation with the Windows operating system. Special emphasis is on a high-fidelity representation of the operation area and a natural sensation of virtual objects using haptic input devices with possibilities for real-time interaction. An important design feature was to provide a habitual laparoscopic environment to the surgeons.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 3: List of the VR applications used for the training of health professionals (continued)

<table>
<thead>
<tr>
<th>Type of intervention</th>
<th>Application title (Platforms)</th>
<th>Author</th>
<th>Studies that have used it</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training of health professionals</td>
<td>GI-Mentor II simulator</td>
<td>Simbionix USA Corporation, Cleveland, Ohio, USA</td>
<td>Cohen et al. (2006) Eversbusch et al. (2004) Shirai et al. (2008)</td>
<td>This is a virtual reality endoscopic trainer. Endoscopic procedures on the GI Mentor are performed on a human-sized mannequin with a modified Pentax ECS-3840F endoscope. The software package allows the user to practice EGD, colonoscopy, and endoscopic retrograde cholangiopancreatography. It allows practice of both technical manoeuvres and cognitive recognition skills. Special hand-eye coordination exercises are incorporated and it provides visual and audible feedback on discomfort.</td>
</tr>
<tr>
<td>Assessment of skills</td>
<td>Minimally Invasive Surgical Trainer - Virtual Reality (MIST-VR)</td>
<td>Personal computer: (Matrox Graphics Inc., 1055 St Regis Blvd., Dorval, Quebec, Canada) Virtual Laparoscopic Interface frame set: (Immersion Corporation, 801 Fox Lane, San Jose, CA, USA)</td>
<td>Gallagher et al. (2004)</td>
<td>MIST stimulator based on a 200-MHz Pentium personal computer (PC) running Windows 95 with 32-MbRAM and a Matrox Mystique 4-MB video card. The laparoscopic interface was a standard Virtual Laparoscopic Interface frame set. It lacks haptic feedback. Simulator tasks enable the acquisition of psychomotor skill rather than cognitive knowledge. Thus this stimulator can teach the basic skills required for all forms of minimally invasive surgery.</td>
</tr>
</tbody>
</table>
### Table 3: List of the VR applications used for the training of health professionals (continued)

<table>
<thead>
<tr>
<th>Type of intervention</th>
<th>Application title (Platforms)</th>
<th>Author</th>
<th>Studies that have used it</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training of health professionals</td>
<td>Minimally Invasive Surgical Trainer - Virtual Reality (MIST-VR)</td>
<td>Mentice Medical Simulation, Gothenburg, Sweden</td>
<td>Grantcharov et al. (2004) Aggarwal et al (2004)</td>
<td>The MIST VR system is run on a desktop PC (400-MHz Pentium II, 64-Mb RAM) and video subsystem employed Matrox Mystique, 8-MB SDRAM video card. The laparoscopic interface input device consisted of two laparoscopic instruments at a comfortable surgical height relative to the operator. Targets appear randomly on the screen and are ‘grasped’ or ‘manipulated’ with performance measured by time, error rate and economy of movement for each hand. Trainees are guided through a series of exercises of progressive complexity. Simulator has tasks that are abstract in nature, enabling the acquisition of psychomotor skill rather than cognitive knowledge.</td>
</tr>
<tr>
<td>Assessment of skills</td>
<td>Minimally Invasive Surgical Trainer, Virtual Reality (MISTVR)</td>
<td>Virtual Presence Ltd., London, SE1 2NL, UK</td>
<td>Grantcharov et al. (2001) Torkington et al. (2001)</td>
<td>Simulates a range of laparoscopic tasks. The system is based on a PC and configured with a Pentium 200-MHz processor, 32 MB of RAM, a 1.6-GB hard drive, a Matrox Mystique 4-MB video card, and a 17-in monitor. It is linked to a frame containing two laparoscopic instruments and a diathermy pedal.</td>
</tr>
</tbody>
</table>
### Table 3: List of the VR applications used for the training of health professionals (continued)

<table>
<thead>
<tr>
<th>Type of intervention</th>
<th>Application title (Platforms)</th>
<th>Author</th>
<th>Studies that have used it</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training of health professionals</td>
<td>AccuTouch Flexible Bronchoscopy Simulator</td>
<td>Immersion Medical, Gaithersburg, MD</td>
<td>Goldmann et al. (2006), Rowe &amp; Cohen (2002)</td>
<td>This VR partial-task trainer consisting of a proxy flexible bronchoscope, a robotic interface, a computer, a monitor, and comprehensive simulation software capable of delivering different realistic bronchoscopy scenarios.</td>
</tr>
<tr>
<td>Assessment of skills</td>
<td>Procedicus arthroscopy simulator</td>
<td>Mentice Corp, Göteborg, Sweden</td>
<td>Gomoll et al. (2008)</td>
<td>Simulator has realistic arthroscopic user interface mimicking the handles of an angled scope and a hook probe. It also has high quality haptic feedback when interacting with the anatomical joint structures.</td>
</tr>
<tr>
<td>Assessment of skills</td>
<td>LapSim surgical trainer</td>
<td>Surgical Science Ltd., Gothenburg/Sweden</td>
<td>Hassan et al. (2006), Langelotz et al. (2005), Larsen et al. (2009), Aggarwal et al (2004), Aggarwal and Tully (2006)</td>
<td>This laparoscopic trainer has tasks that are more realistic than those of the MIST-VR, involving structures that are deformable and may bleed. In addition to its basic skills module, has a software module for a laparoscopic cholecystectomy called the LapSim Dissection. Simulator creates a virtual laparoscopic system using a computer (Windows XP), a video monitor, and laparoscopic interface containing two pistol-grip instruments and a diathermy pedal. The LapSim software contains the basic modules referred to as “cutting, clip application, and coordination.” The system used does not possess haptic feedback.</td>
</tr>
</tbody>
</table>
Table 3: **List of the VR applications used for the training of health professionals (continued)**

<table>
<thead>
<tr>
<th>Type of intervention</th>
<th>Application title (Platforms)</th>
<th>Author</th>
<th>Studies that have used it</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training of health</td>
<td>LapMentor</td>
<td>Simbionix, Cleveland,</td>
<td>Aggarwal et al (2004)</td>
<td>This stimulator enables the trainee to perform a complete laparoscopic cholecystectomy with the benefit of force feedback. It has augmented their basic skills programmes to incorporate parts of real procedures, allowing trainees to learn techniques they would use in the operating theatre.</td>
</tr>
<tr>
<td>professionals</td>
<td></td>
<td>Ohio, USA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment of skills</td>
<td>UroMentor</td>
<td>Simbionix Ltd, Israel</td>
<td>Knoll et al. (2005)</td>
<td>This simulator provides a realistic simulation of rigid and flexible ureterorenoscopy (URS).</td>
</tr>
<tr>
<td>Assessment of skills</td>
<td>Xitact LS500 laparoscopy</td>
<td>Xitact SA, Morges,</td>
<td>Schijven et al. (2005)</td>
<td>This simulator comprises tasks such as dissection, clip application and tissue separation, the integration of which can produce a procedural trainer. An independent Endoscope is provided, which allows cont rolling an endoscope with angled optics. The brain of the system is a single PC machine. Simulation software include 4 sets of training modules: camera navigation, clip and cut training, superficial dissection, and advanced dissection. It differs from theMIST-VR and LapSim in that it incorporates a physical object, the ‘virtual abdomen’, with force feedback.</td>
</tr>
</tbody>
</table>
Table 3: **List of the VR applications used for the training of health professionals (continued)**

<table>
<thead>
<tr>
<th>Type of intervention</th>
<th>Application title (Platforms)</th>
<th>Author</th>
<th>Studies that have used it</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment of skills</td>
<td>SIMENDO VR simulator</td>
<td>Delta Tech, Delft, the Netherlands</td>
<td>Verdaasdonk et al. (2008)</td>
<td>Simulator designed to train eye–hand coordination skills using abstract tasks such as camera navigation and basic drills such as pick-and-place tasks</td>
</tr>
</tbody>
</table>
Types of Simulators

As computer technology has improved in performance and cost-efficiency, so to have the simulator systems. Simulation-based training has gained significant momentum and wide variety of simulators are currently available for training of health professionals in surgery ranging from inanimate video trainers, human patient simulators, to more recently virtual reality (VR) computer-based trainers (Rehrig el al 2008).

A total of 12 applications (platforms) were identified in this review (Table 3). The features of current VR surgical simulators enable their use to train surgical skills and/or to make inferences about levels of surgical performance. In their latter role, VR training systems were designed to measure various aspects of performance, such as motion and efficiency characteristics, errors, and time to complete a specified task. The validation of these measurement capabilities has been the subject to numerous investigations, with the aim of ultimately establishing that training in a VR environment improves clinical performance. Additionally, a number of studies have been undertaken to demonstrate that skills acquired during VR training transfer to the operating room.

In this review, the studies appraised were classified into nine themes based on the purpose or main aim of the studies (Table 4)
Table 4: **Key themes identified as a result of appraisal of articles**

<table>
<thead>
<tr>
<th>Theme</th>
<th>Description of the theme</th>
<th>Studies included</th>
<th>General conclusion</th>
</tr>
</thead>
</table>
Table 4: **Key themes identified as a result of appraisal of articles (continued)**

<table>
<thead>
<tr>
<th>Theme</th>
<th>Description of the theme</th>
<th>Studies included</th>
<th>General conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>VR shortens the learning curve</td>
<td>VR simulation technologies to shorten the learning curve for achievement of proficiency</td>
<td>Aggarwal and Black (2006) Gallagher et al. (2004) Aggarwal et al., (2004)</td>
<td>VR simulation has the ability to discriminate the performance between experts and novices. It may be useful in identifying subset of novices who has difficulty acquiring the psychomotor skills.</td>
</tr>
</tbody>
</table>
Table 4: Key themes identified as a result of appraisal of articles (continued)

<table>
<thead>
<tr>
<th>Theme</th>
<th>Description of the theme</th>
<th>Studies included</th>
<th>General conclusion</th>
</tr>
</thead>
</table>
### Table 4: Key themes identified as a result of appraisal of articles (continued)

<table>
<thead>
<tr>
<th>Theme</th>
<th>Description of the theme</th>
<th>Studies included</th>
<th>General conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjunct to traditional methods</td>
<td>Virtual reality simulation is seen as an adjunct to traditional methods of training, and not as an alternative.</td>
<td>Aggarwal et al. (2004) Botden et al. (2008) Cakmak et al. (2005) Gurusamy et al. (2008)</td>
<td>Simulators currently have the ability to teach basic surgical skills, enabling novice surgeons to progress along the early part of the learning curve before entering the operating theatre. However, surgeons will still need to reach expert levels of skill in the operating theatre; further training to achieve this is essential.</td>
</tr>
<tr>
<td>Assessment of psychomotor performance</td>
<td>VR simulators as a means of training and objective assessment of psychomotor performance</td>
<td>Gallagher et al. (2004) Grantcharov et al. (2004) Hassan et al. (2006)</td>
<td>VR systems allow repeated practice of standardized tasks, and provide unbiased and objective measurements of laparoscopic performance. Surgeons who received VR simulator training showed significantly greater improvement in performance in the OR than those in the control group. VR surgical simulation is therefore a valid tool for training of laparoscopic psychomotor skills and could be incorporated into surgical training programmes.</td>
</tr>
<tr>
<td>Transfer of skills</td>
<td>Virtual reality skills can translate into improved outcome in the ‘real’ procedures e.g. operating room</td>
<td>Grantcharov et al. (2004) Aggarwal et al., (2004)</td>
<td>Surgeons who received VR simulator training showed significantly greater improvement in performance in the OR than those in the control group.</td>
</tr>
<tr>
<td>Theme</td>
<td>Description of the theme</td>
<td>Studies included</td>
<td>General conclusion</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Improvement of surgical skills</td>
<td>Virtual reality training can lead to improvement of decision-making and procedure-specific surgical skills.</td>
<td>Aggarwal and Tully (2006)</td>
<td>Surgeon with minimal laparoscopic experience can improve their skills during short-phase training on a VR procedural module. Thus, VR simulation may be useful for the early part of the learning curve for surgeons.</td>
</tr>
<tr>
<td>Importance of haptic feedback</td>
<td>Botden et al. (2008) Gurusamy et al (2008)</td>
<td></td>
<td>Haptic feedback is a crucial sensorial modality in virtual reality interactions. This study focuses on the important of haptic feedback in laparoscopic suturing training and the additive value of virtual reality simulation.</td>
</tr>
</tbody>
</table>
Emergent themes from the studies appraised

In this review, the studies appraised were classified into nine themes based on the purpose or main aim of the studies. These themes were as follows:

1. Virtual Reality training improves performance tasks:

Ten studies looked at the effectiveness of VR training on performance skill improvement of trainees. The type of studies included; one meta-analysis (Gurusamy et al., 2008), one systematic review (Aggarwal et al., 2004), six Randomized Controlled Trials (Cohen et al., 2006; Grantcharov et al., 2004; Larsen et al., 2009; Rowe & Cohen, 2002; Shirai et al., 2008; Verdaasdonk et al., 2008), one before-after intervention study (Hassan et al., 2006) and two Quasi-experimental study (Aggarwal & Black, 2006; Aggarwal and Tully, 2006). Overall, results of these studies demonstrated that the VR training groups performed the tasks more quickly and with less error than the groups without training. In spite of the different models used, and the lack of similarity between the training and evaluation tasks, the VR group performed consistently better in one or more of the measured parameters. It has also shown that even in people who are at the beginning of their surgical careers, virtual reality training reduces the operating time and increases the procedure accuracy. Thus, there is convincing evidence that VR training is a useful training method especially for novices with limited experience.

The meta-analysis by Gurusamy et al., (2008) showed that of the 23 trials included 12 trials compared VR and no training or standard laparoscopic training (SLT). The results showed that for trainees with no laparoscopic experience VR group performed the tasks more quickly than the group without training (standardized mean difference − 1.09, 95 per cent confidence
interval – 1.50 to 0.68). Results also showed a lower error score in the VR group. However, this difference was not statistically significant. The error score was assessed by dropping the object, perforation or tear of the object, slack ligature, or by composite error score from the computer. Trials which compared virtual reality versus standard laparoscopic training reported noted a statistically significant shorter operative time and lower error score in the VR group than the SLT group.

The systematic review by (Aggarwal et al., 2004) demonstrated that VR simulators have the ability to teach basic laparoscopic skills, enabling novice surgeons to progress along the early part of the learning curve. Randomized trial by Grantcharov et al., (2004) found that surgeons who received VR simulator training showed significantly greater improvement in performance in the operative room than those without VR training. Before-after intervention study (Hassan et al., 2006) concluded that a three-day practical course for laparoscopic surgery improved laparoscopic skills of residents. However, they also noted that advanced residents benefit most from the course.

One study (Cohen et al., 2006) showed that VR simulation does not have a significant additional value in laparoscopic suturing training, over traditional box trainers. They found that while an important advantage of VR is objective assessment of the performance, the participants of the study preferred the traditional box trainer over the VR simulator.

2. Effectiveness as a training tool

Quasi-experimental study by Aggarwal & Black, (2006) looked at the effectiveness of VR simulator as training tool. The reviewers noted that over the six sessions with VR simulator, the inexperienced group made significant improvements in performance for time taken (p = 0.007) and contrast fluid usage (p = 0.021), achieving similar scores at the end of the training
program to the experienced group. This study has shown a vascular interventional virtual reality simulator to be a valid tool for both assessment and training in endovascular skills.

The intention of Quasi-experimental study by Aggarwal and Tully (2006) was to evaluate a VR ectopic module (LAPSIM), in terms of its validity as a training and assessment tool for trainee gynaecological surgeons. There were significant differences between the gynaecological surgeons of three levels of difference at the second repetition of the ectopic module for time taken (median 551.1 versus 401.2 versus 249.2 seconds, P = 0.001), total blood loss (median 304.2 versus 187.4 versus 123.3 ml, P = 0.031) and total instrument path length (median 17.8 versus 8.3 versus 6.8 m, P = 0.023). Investigators concluded that Gynaecological surgeons with minimal laparoscopic experience can improve their skills during short-phase training on a VR procedural module.

3. Learning curves on simulators:

Three studies analysed the learning curves of experts and novices. Systematic review by Aggarwal et al., (2004) noted that VR simulators have the ability to teach basic laparoscopic skills, enabling novice surgeons to progress along the early part of the learning curve before entering the operating theatre. Authors found that subjects improved their total time taken and accuracy of movement throughout the sessions with VR simulators.

Three Quasi-experimental studies (Aggarwal and Black, 2006; Gallagher et al., 2004; Aggarwal and Tully, 2006) looked at whether training on the simulator could lead to an improvement in the skills of inexperienced operators. Results from both these studies indicated that trainees with minimal experience can improve their performance tasks during short-phase training on a VR simulator. The study by Gallagher et al. (2004) demonstrated that after three trials with MIST-VR system the mean of the medical students’ performance is
not significantly different from the mean of the experienced group of surgeons with respect to error, economy of movement of the right and left hands, or economy of diathermy. They also noted that statistically significant difference in total time (p = 0.0058). Investigators concluded that the MIST-VR system is capable of measuring the variability in performance between experts and novices. The study by Aggarwal and Black, (2006) found that over the six sessions VR simulator, inexperienced group made significant improvements in performance for time taken (p = 0.007) and contrast fluid usage (p = 0.021), achieving similar scores at the end of the training program to the experienced group. Authors concluded that VR simulation may be useful for the early part of the learning curve for surgeons who wish to expand their surgical interests. The results of the study by Aggarwal and Tully (2006) showed that Gynaecological surgeons with minimal laparoscopic experience can improve their skills during short-phase training on a VR procedural f t module. In contrast, experienced operators showed non-significant improvements. Authors concluded that VR simulation may be useful for the early part of the learning curve for surgeons who wish to learn to perform laparoscopic salpingectomy for ectopic pregnancy.

4. Virtual Reality simulators as assessment devices:

Eleven studies looked at the feasibility and validity of performance assessments with virtual-reality simulators. The type of studies included; one meta-analysis (Gurusamy et al., 2008), one systematic review Aggarwal et al., (2004), four Randomized Controlled Trials (Botden et al., 2008; Grantcharov et al., 2004; Cohen et al., 2006; Eversbusch et al., 2004), one narrative review (Cakmak et al., 2005), two Quasi-experimental studies (Gallagher et al., 2004; Aggarwal & Tully, 2006) and two before and after studies (Grantcharov et al., 2001; Hassan et al., 2006). Studies to confirm the role of virtual reality simulators as assessment devices
have mainly concentrated on the demonstration of construct validity, with experienced surgeons completing the tasks simulators significantly faster, with lower error rates and greater economy of movement scores.

5. Virtual Reality simulators adjunct to traditional methods

Four studies looked at the effectiveness of Virtual Reality simulators as a supplement to standard training. The results of the meta-analysis by Gurusamy et al., (2008) showed that there is convincing evidence for VR training as a useful supplement to standard laparoscopic training in surgical residents with limited laparoscopic experience. The systematic review by Aggarwal et al., (2004) noted that virtual reality simulation should be seen as an adjunct to traditional methods of training, and not as an alternative.

6. Assessment of psychomotor performance

Three studies looked at the role of the VR simulators as an assessment of psychomotor performance (Gallagher et al., 2004; Grantcharov et al., 2004, Hassan et al., 2006). One Quasi-experimental study (Gallagher et al., 2004) demonstrated that the Minimally Invasive Surgical Trainer—Virtual Reality (MIST-VR) system is capable of evaluating the psychomotor skills necessary in laparoscopic surgery and discriminating between experts and novices. They also noted that when compared to established performance criteria, the subjects can be stratified according to psychomotor ability.

One randomized trial (Grantcharov et al., 2004) concluded that VR surgical simulation is a valid tool for training of laparoscopic psychomotor skills and could be incorporated into surgical training programmes.
7. Evidence for the role of virtual reality simulators in the transfer of skills:

Before the incorporation of simulators into the general surgical curriculum, it is necessary to show transfer of surgical skills to real operations. The systematic review by Aggarwal et al., (2004) noted that VR trained skills may be transferable and lead to improved performance intra-operatively compared to standard training. However, they identified that there is a weak evidence for the role of virtual reality simulators in the transfer of skills.

Training within a skills laboratory cannot fully mimic the clinical setting. Consequently, students may encounter problems such as increased responsibility, inadequate supervision, and difference in equipment use when they attempt to transfer the clinical skills they learned in the laboratory to the clinical setting. Thus, more work is needed to explore these problems encountered by students performing clinical skills learned in a skills laboratory.

8. Improvement of surgical skills

Skills in surgery can be increased in a clinically relevant manner using proficiency based virtual reality simulator training. Gurusamy et al., 2008 showed that in people with nolaparoscopic experience, VR training is better than no training in relation to the time taken to perform a task, improving accuracy and decreasing error. Various methods of training were used in the trials but the benefits were consistent. For young surgeons at the beginning of their laparoscopic training, VR training reduced the operating time, error and unnecessary movements during laparoscopic cholecystectomy. The authors concluded that there is convincing evidence for VR training to be considered as a useful supplement to standard laparoscopic training (SLT). Aggarwal and Tully (2006) looked at the effectiveness of VR
ectopic module (LAPSIM) to be integrated into a skills curriculum. There were significant differences between the groups at the second repetition of the ectopic module for time taken (median 551.1 versus 401.2 versus 249.2 seconds, \( P = 0.001 \)), total blood loss (median 304.2 versus 187.4 versus 123.3 ml, \( P = 0.031 \)) and total instrument path length (median 17.8 versus 8.3 versus 6.8 m, \( P = 0.023 \)). The learning curves for intermediate and novice groups were steeper and longer when compared with the experts. However, inexperienced subjects managed to achieve similar levels of skill towards the end of the training period. Authors concluded that Gynaecological surgeons with minimal laparoscopic experience can improve their skills during short-phase training on a VR procedural module.

9. VR training with haptic feedback (touch and pressure feedback):

Two studies looked at the effectiveness of haptic feedback (Botden et al., 2008; Gurusamy et al., 2008). Meta-analysis by Gurusamy et al., (2008) reported that improving the fidelity by haptic feedback may increase trainee satisfaction and the enthusiasm to learn on VR models. The findings of this review suggest that haptic feedback should be an important component for inclusion into future simulator technology. Improving the reliability by haptic feedback may increase trainee satisfaction and the enthusiasm to learn on VR models. Randomized Controlled Trial by Botden et al., (2008) identified that VR simulations with haptic feedback imply better training effects and a better transfer of skills to the clinical setting.
Discussion

The purpose of this review was to assess the effectiveness of immersive Virtual Reality se training applications for improvement of clinical expertise of health professionals. The review included a methodical search of all the relevant medical, educational and computer literature databases, and included randomized trials that evaluated the effectiveness of VR training applications for improvement of clinical skills of health professionals. The search strategy identified 1379 studies. Stage one resulted in 368 eligible studies. Stage two resulted in 24 eligible studies to be included in the review. These eligible studies were based on a total of 1174 participants and reported the results of one meta-analysis, one systematic review, nine Randomised Control Trials (RCTs) and 13 other studies of different designs.

A total of twelve VR simulators were found to be evaluated in the different studies. These included Vascular Interventional Surgical Trainer (VIST simulator), Surgical Education Platform (SEP) trainer, VSOOne (Virtual Endoscopic Surgery Training (VEST) system), GI-Mentor II simulator, Minimally Invasive Surgical Trainer - Virtual Reality (MIST-VR), AccuTouch Flexible Bronchoscopy Simulator, Procedicus arthroscopy simulator, LapSim surgical trainer, LapMentor, UroMentor, Xitact LS500 laparoscopy simulator, SIMENDO VR simulator. VR training systems were designed to measure various aspects of performance, such as motion and efficiency characteristics, errors, and time to complete a specified task. The majority of these applications were found to be effective, in as much as they were found to significantly improve the desired clinical skills of health professionals.

The different reviews and clinical trials and other studies evaluated five different outcomes to assess the effectiveness of different VR simulation applications. The different outcomes include, performance task on the simulator (time to complete the procedure, error score,
economy of movement score), assessment of the learning curves, skill transferability, reaching benchmark criteria, patient discomfort and satisfaction.

In this review, the studies appraised were classified into nine overlapping ‘themes’ based on the purpose or main aim of the studies. These themes were as follows: Virtual Reality training improves skills, effectiveness as a training tool, learning curves on simulators, Virtual Reality simulators as assessment devices, Virtual Reality simulators adjunct to traditional methods, assessment of psychomotor performance, evidence for the role of virtual reality simulators in the transfer of skills, improvement of surgical skills, and VR training with haptic feedback.

Critical appraisal of the available evidence identified several advantages of the VR training environments in health professionals and provides excellent direction for medical educators who are trying to make training decisions. In brief, these are as follows:

First, there was evidence that VR training improves skills and is an effective training method especially for novices with limited experience. Trainees who received VR simulator training showed significantly greater improvement in time, error and economy of movement scores in the operating room than those with no training. Evidence in support of this principle is based on meta-analysis by Gurusamy et al (2008) and several RCTs. In the meta-analysis by Gurusamy et al (2008) there is convincing evidence for VR training to become useful supplement to standard laparoscopic training in laparoscopic cholecystectomy in surgical residents with limited laparoscopic experience.

Cohen et al. (2006) reported that surgical fellows who underwent GI Mentor training performed significantly better during the early phase of real colonoscopy training. In the RCT by Grantcharov et al. (2004), surgeons who received VR simulator training showed
significantly greater improvement in performance in the operating room than those in the control group and therefore, VR simulator training could be incorporated into surgical training programmes. Larsen et al. (2009) commented that skills in laparoscopic surgery can be increased in a clinically relevant manner using proficiency based virtual reality simulator training.

Various methods of training were used in these trials but the benefits were consistent. It has been also shown that even in people who are at the beginning of their surgical careers; virtual reality training reduces the operating time and increases the procedure accuracy. Cakmak et al. (2005) introduced VSOne as fully-featured virtual reality training system for laparoscopic surgery. They found that for basic skills training VSOne is suited as well as conventional training, for complex surgical procedure training, VR based training is superior due to the outstanding benefits like the 24 hour availability and reduced training costs.

Before the incorporation of VR simulators into the training programmes, it is necessary to show transfer of skill to real operations. Virtual Reality simulators are likely to be successful in the transfer of skills. Grantcharov et al. (2004) reported in their review VR trained skills may be transferable and lead to improved performance intra-operatively compared to standard training.

Second, virtual reality simulator model seems to be a promising objective method for assessment of clinical skills. It has been demonstrated in a number of studies that the VR simulator is a valid tool for the assessment of clinical skills. Eversbusch et al. (2004) found that the GI Mentor is a valid tool for the evaluation of psychomotor skills in gastrointestinal endoscopy. Gallagher, et al., 2004 reported that the MIST-VR system can measure clinical skills necessary in laparoscopic surgery, as well as the variability in performance between subjects with similar experience.
The study by Cohen et al. (2006) found that both objective and subjective assessments of competency were significantly improved in the simulator group by using the mixed-effects model. Studies to confirm the role of virtual reality simulators as assessment devices have mainly concentrated on the demonstration of construct validity, with experienced surgeons completing the tasks simulators significantly faster, with lower error rates and greater economy of movement scores.

VR model can be considered as method for the assessment and evaluation of psychomotor skills and discriminating. In the study by Gallagher et al. (2004), investigators found that MIST-VR system is capable of evaluating the psychomotor skills necessary in laparoscopic surgery and discriminating between experts and novices. Hassan et al. (2006) tested the psychomotor skills acquisition of residents attending a three-day laparoscopic course using virtual reality simulator LapSim and found that residents with some degree of experience in laparoscopic surgery excluding novices profit mostly from laparoscopic skills courses when psychomotor skills are assessed by a virtual reality simulator.

One study focused on the additive value of VR simulation for laparoscopic suturing training over Box method of training how to suture using laparoscopic surgical tools. This study by Botden et al. (2008) found that VR simulation does not have a significant additive value in laparoscopic suturing training, over traditional box trainers.

Third, studies that assessed the ability of virtual reality trainers to teach clinical skills have analysed the learning curves of experts and novices. Evidence in favour of this came from the two group comparison studies by Aggarwal and Black (2006).
Results of this study showed that ability of VR simulation technologies to shorten the learning curve for achievement of proficiency on real cases when compared with traditional training methods. Authors concluded that VR simulation may be useful for the early part of the learning curve for surgeons who wish to expand their endovascular interests.

Fourth, the haptic feedback is an important feature in laparoscopic suturing simulation. This opinion is supported by the outcome of the study by Botden et al. (2008).

The findings in this review need to be interpreted in the light of several of its limitations. Despite best efforts, not all grey literature or unpublished work or work in progress could not traced. Because of this, this review may have missed studies that otherwise be representative of the entire range of interventions, not just those targeted at surgical trainees and enabling improvement of those skills. Possibly these other studies either have less impressive results, or they might be evaluating new technologies that have a hard time in finding their way to peer reviewed literature, and then again, present uninteresting findings and therefore often are lost.

At present, Virtual Reality applications are deployed widely throughout surgical training programs and serve as the primary platform for laparoscopic skills training. There is still limited evidence coming from control studies to support its position in other disciplines of medicine (e.g. emergency medicine, neurology and specific clinical problems). However, currently there is no agreement regarding the optimal VR assessment tool for training of health professionals. There was no comparison between different VR models in the review studies. Thus, further research work is required to make head-to-head comparisons to find out best suited VR application for each discipline in medicine.
Most of these studies did not make a direct comparison between the actual and virtual environment, but rather based the comparison on simple parameters, such as time to completion, to evaluate the effectiveness of such environments. Only very few recent trials have used predefined criteria or ‘benchmark levels’ for assessment of the effectiveness of virtual reality training.

Training within VR simulations cannot fully mimic the clinical setting. Consequently, students may encounter problems such as increased responsibility, inadequate supervision, and difference in equipment use when they attempt to transfer the clinical skills they learned from VR simulator environment to the clinical setting. There is a lack of well designed studies addressing the important issues of transferability, problems encountered by students during transferability and retention of skills over time. Thus, more work is needed to explore these problems encountered by students performing clinical skills learned in a skills laboratory.

Many studies of the effectiveness of simulators often are limited in that they measure performance using the same training simulator, which may favour those who have trained on the simulator itself. In other words, apparently improved performance may not translate to actual patient care. Of the studies that have extended laboratory simulation to health professional training, none have addressed the technological challenges relating to use of VR applications in medicine. Some of the simulators -described in the literature are still in development phase and not have been fully validated in adequate trials. These technological challenges should be evaluated in further research.

A key to the success of simulation training is integrating it into traditional education programs. Training in a simulated environment is a new, additional step in the learning
process — a step between classroom instruction and actual clinical instruction with real patients. Only few trials assessed VR training as part of a surgical training curriculum. Hence further trials are necessary to assess the potential role, duration and scope of such training.

Another requirement for integrating VR into educational curriculum is safety and system usability. Though the latest VR tools seem to have minor or no side effects, future research required to confirm these results. Thus, further research is required, both on technological side and on VR issues such as transfer of learning, optimal VR application, safety and the psychological and social impact of the technology use, if medical skills laboratories are to remain an integral component of medical education.

Eventhough the review becomes limited in scope, the justification for excluding studies carried out with medical student participants is presumably because the objective was to identify VR tools that are effective for improving clinical skills in health professionals. Medical students are not in a position to make independent decisions at least in the contexts where are they work, and therefore synthesis of evaluations that included medical students would not add anything to the intended audience (i.e. post certification health professionals). Thus the pedagogical inference becomes somewhat limited and review findings would only be applied in the context of post certification health professionals. In addition to these limitations, there were several other limitations of the component studies including; small participant numbers and failure to blind patient and clinician to the nature of the intervention.

In summary, a broad review of the literature was conducted to assess the effectiveness of virtual reality based immersive training applications for improvement of clinical expertise of health professionals. Literature data bases were searched with specified search terms and resulted in retrieval of 25 studies. This review is thus based on one meta-analysis, one
systematic review and the remaining 23 critically appraised RCTs and other different study
designs. Critical appraisal of these studies resulted in identification of several different types
of Virtual Reality applications, and different types of outcomes related to Virtual Reality
simulations and clinical skill training.

In general, the following key features emerged from this review and presented below:

1. Virtual Reality simulators can be considered a useful tool for improvement of clinical
   skills performance of health professionals and thereby shorten the learning curve for
   achievement of proficiency on real cases.

2. VR simulators allow assessment of clinical skills and discriminating between experts
   and novices.

3. Haptic feedback should be an important component of VR simulators.

4. Limited convincing evidence to support role of VR simulators in disciplines of
   medicine other than surgical specialties.

5. Lack of well designed studies addressing the important issues such as optimal VR
   application, problems encountered during transferability, safety, psychological and
   social impact of the technology use and retention of skills over time.

This review had few limitations that need to be taken into consideration as the results of this
review are interpreted. These were incomplete search for the gray literature and non-
availability of full texts of specific studies that nevertheless passed the initial first pass
criteria. Most of the randomised controlled trials included in the review are small, sometimes
with as few as 10 participants in each arm of the trial.
It is important to consider that the small numbers involved mean that the trials can only detect large differences in outcome between the groups at baseline, controlling for confounding. In addition, the evidence considered in this review displayed some limitations including biases and failure to assess outcomes in a manner that is blind to intervention task that may impact on the validity of individual studies.
Conclusion

Virtual Reality applications have been found to improve procedural skills compared with standard or no training at all and it is likely that their role in training of medical personnel will grow. VR applications, in combination with existing opportunities to work with real patients, can increase the range of experience to learn about and deal with medical problems as learners and practitioners, ensure consistency of training experiences, and improve the acquisition of clinical skills. In order to be effective, however, such activity needs to be part of a broader picture, supporting and work closely with actual clinical practice. In general most of the findings from this review demonstrate that, VR simulators can be considered a useful tool for improvement of clinical skills performance of health care professionals.

Healthcare's potential use of interactive 3D technologies is broad. To date, most of the media's attention has centered on two application areas: surgical training and planning, and computer-aided surgery systems. However, the possible uses are much broader.

Recent evidence on the effectiveness of using virtual reality training applications for improvement of clinical skills of health professionals is limited but sufficiently encouraging to justify additional clinical trials in this area. In spite of these successes, there remain several gaps in the literature regarding the effectiveness of using VR training environments. Further research is required addressing these important issues if VR simulations are to remain an integral component of medical education. In addition, significant efforts are still required to move VR into economical success and therefore routine use in medical training. Well designed studies to improve our understanding of their effects on training will allow them to be used more intelligently to improve performance of health professionals, reduce their errors and ultimately, promote patient safety. Although such studies will be difficult and costly, they may be justified to determine how this technology can best be applied.
References


37. Evidence-Based Answers to Clinical Questions for Busy Clinicians. (2006), The Centre for Clinical Effectiveness, Monash Institute of Health Services Research, Melbourne, Australia


Glossary

**Bias**  Deviation of results or inferences from the truth, or processes leading to such deviation. Any trend in the collection, analysis, interpretation, publication, or review of data that can lead to conclusions that are systematically different from the truth.

**Blinding/blinded**  A trial is fully blinded if all the people involved are unaware of the treatment group to which trial participants are allocated until after the interpretation of results.

**Box (video) trainer**  incorporates conventional laparoscopic equipment, is a relatively inexpensive and highly versatile device that enables the training on animal parts as well as synthetic inanimate models.

**Case control study**  An epidemiological study involving the observation of cases (persons with disease, such as cervical cancer) and a suitable control (comparison, reference) group of persons without the disease. The relationship of an attribute to the disease is examined by comparing retrospectively the past history of the people in the two groups with regard to how frequently the attribute is present.

**Case series**  Analysis of series of people with the disease (there is no comparison group in case series).

**Cohort study**  The analytic method of epidemiologic study in which subsets of a defined population can be identified who are, have been, or in the future may be exposed or not exposed in different degrees, to a factor or factors hypothesised to influence the probability of occurrence of a given disease or other outcome. Studies usually involve the observation of a large population, for a prolonged period (years), or both.
**Confidence interval**  The computed interval with a given probability, e.g. 95%, that the true value of a variable such as mean, proportion, or rate is contained within the interval. The 95% CI is the range of values in which it is 95% certain that the true value lies for the whole population.

**Confounder**  A third variable that indirectly distorts the relationship between two other variables, because it is independently associated with each of the variables.

**Confounding**  A situation in which the measure of the effect of an exposure on risk is distorted because of the association of exposure with other factor(s) that influence the outcome under study.

**Controls**  In a randomised controlled trial (RCT), controls refer to the participants in its comparison group. They are allocated either to placebo, no treatment, or a standard treatment.

**Descriptive study**  A study concerned with, and designed only to describe the existing distribution of variables, without regard to causal or other hypotheses.

**Effectiveness**  A measure of the extent to which a specific intervention, procedure, regimen, or service, when deployed in the field in routine circumstances, does what it is intended to do for a specific population.

**Evidence based**  Based on valid empirical information

**Experimental study**  A study in which the investigator studies the effect of intentionally altering one or more factors under controlled conditions

**Grey literature**  That which is produced by all levels of government, academics, business and industry, in print and electronic formats, but which is not controlled by commercial publishers.
**Incidence**  The number of new events (cases, e.g. of disease) occurring during a certain period, in a specified population.

**Inclusion/exclusions**  We use validated search and appraisal criteria to exclude unsuitable papers. Authors are then sent exclusion forms to provide reasons why further papers are excluded.

**LAP Mentor**  multi-disciplinary LAP surgery simulator enables simultaneous hands-on practice for a single trainee or a team. The system offers training opportunities to new and experienced surgeons for everything from perfecting basic laparoscopic skills to performing complete laparoscopic surgical procedures.

**LapSim**  laparoscopic trainer has tasks that are more realistic than those of the MIST-VR involving structures that are deformable and may bleed.

**Meta-analysis**  The process of using statistical methods to combine the results of different studies. The systematic and organised evaluation of a problem, using information from a number of independent studies of the problem.

**Minimally Invasive Surgery Trainer-Virtual reality (MIST-VR)**  Computer-based laparoscopic simulator comprises two standard laparoscopic instruments held together on a frame with position-sensing gimbals. These are linked to a Pentium personal computer and movements of the instruments are relayed in real time to a computer monitor. Targets appear randomly on the screen and are ‘grasped’ or ‘manipulated’, with performance measured by time, error rate and economy of movement for each hand.

**Non-systematic review**  A review or meta-analysis that either did not perform a comprehensive search of the literature and contains only a selection of studies on a clinical question, or did not state its methods for searching and appraising the studies it contains.
**Odds ratio (OR)**  A measure of the degree or strength of an association. In a case control or a cross-sectional study, it is measured as the ratio of the odds of exposure (or disease) among the cases to that among the controls.

**PICO**  Population, intervention, comparison, and outcome. The current reporting requirements of systematic reviews are: how many RCTs, how many participants in each, comparing what with what, in what type of people, with what results.

**Power**  A study has adequate power if it can reliably detect a clinically important difference (i.e. between two treatments) if one actually exists. The power of a study is increased when it includes more events or when its measurement of outcomes is more precise.

**Prevalence**  The proportion of people with a finding or disease in a given population at a given time.

**P value**  The probability that an observed or greater difference occurred by chance, if it is assumed that there is in fact no real difference between the effects of the interventions. If probability is \(< \frac{1}{20}\) (which is when the P value is less than 0.05), then the result is conventionally regarded as being "statistically significant".

**Randomised**  We aim to provide an explanation of how a trial is quasi-randomised in the *Comment* section.

**Randomised controlled trial**  An epidemiologic experiment in which subjects in a population are randomly allocated into groups to receive or not receive an experimental preventive or therapeutic procedure, manoeuvre, or intervention. Randomised controlled trials are generally regarded as the most scientifically rigorous method of hypothesis testing available in epidemiology.
**Relative risk (RR)**  The ratio of the risk of disease or death among the exposed to the risk among the unexposed. It is a measure of the strength or degree of association applicable to cohort studies and RCTs.

**Systematic review**  Literature review reporting a systematic method to search for, identify and appraise a number of independent studies.

**Validity**  The soundness or rigour of a study. A study is internally valid if the way it is designed and carried out means that the results are unbiased and it gives you an accurate estimate of the effect that is being measured. A study is externally valid if its results are applicable to people encountered in regular clinical practice.
Appendix A: Included Studies


Appendix B: Excluded studies annotated by reasons for exclusion


Almeida, S., L. Brasil, et al. (2007), The diagnosis support system for ischemic cardiopathy: a case study in the context of IACVIRTUAL project, Springer.Excluded: Not related to education of health professionals


Excluded: Not related to education of health professionals


Cyberpsychology & behavior 9(2): 129-32.
Excluded: Not related to education of health professionals


Studies in health technology and informatics 85: 144-9, Excluded: Not related to education of health professionals


Flannery, K. A. and R. Walles (2003), "How does schema theory apply to real versus virtual memories?” Cyberpsychology & behavior 6(2): 51-9, Excluded: Not related to education of health professionals


Ganger, P. (2005), "Future Surgical Training"
Complications in surgery: 9, Excluded: Abstract not available


Biological psychiatry 60(7): 752-9. Excluded: Not related to education of health professionals ?


Kaltenborn, K. F. and O. Rienhoff (1993), "Virtual reality in medicine" Methods of information in medicine 32(5): 407-17, Excluded: Published before 1999


Kayser, K., J. Görtler, et al. (2008). "Image standards in Tissue-Based Diagnosis(Diagnostic Surgical Pathology)." Diagnostic Pathology 3(1): 17, Excluded: Abstract not available


Kerner, K. F., C. Imielinska, et al. (2003), "Augmented Reality for teaching endotracheal intubation: MR imaging to create anatomically correct models" AMIA ... Annual Symposium proceedings / AMIA Symposium: 888, Excluded: Not related to education of health professionals


Kim, J. and M. A. Srinivasan (2005), "Characterization of viscoelastic soft tissue properties from in vivo animal experiments and inverse FE parameter estimation" Medical image computing and computer-assisted intervention 8(Pt 2): 599-606, Excluded: Not related to education of health professionals


Lamata, P., E. J. Gomez, et al. (2006). "Tissue consistency perception in laparoscopy to define the level of fidelity in virtual reality simulation" *Surgical endoscopy* 20(9): 1368-75. Excluded: Not related to education of health professionals


Lee, W., H. Kim, et al. (2004), "CT arthrography and virtual arthroscopy in the diagnosis of the anterior cruciate ligament and meniscal abnormalities of the knee joint", Excluded: Abstract not available


Lotan, M., S. Yalon-Chamovitz, et al. (2009), "Improving physical fitness of individuals with intellectual and developmental disability through a Virtual Reality Intervention Program" Research in developmental disabilities **30**(2): 229-39, Excluded: Not related to education of health professionals


Mergner, T., G. Schweigart, et al. (2005), "Human postural responses to motion of real and virtual visual environments under different support base conditions." *Experimental brain research.* Experimentelle Hirnforschung 167(4): 535-56, Excluded: Not related to education of health professionals


Murphy, D., B. Challacombe, et al. (2007), "[Equipment and technology in robotics]." Archivos espanoles de urologia 60(4): 349-55, Excluded: Abstract not available


Sanchez-Vives, M. V. and M. Slater (2005), "From presence to consciousness through virtual reality" Nature reviews 6(4): 332-9, Excluded: Not related to education of health professionals

Sankaranarayanan, G., S. Arikatla, et al. (2009)."Face validation of the virtual basic laparoscopic skill trainer (VBLaST)." Studies in health technology and informatics 142: 286 Excluded: Not related to education of health professionals


Schijven, M. P. and J. J. Jakimowicz (2005), "Validation of virtual reality simulators: Key to the successful integration of a novel teaching technology into minimal access surgery." Minimally invasive therapy & allied technologies 14(4): 244-6, Excluded: Not related to education of health professionals


Schneider, S., A. Hood, et al. (2005), "Virtual reality intervention for chemotherapy symptoms” Oncology Nursing Society, Orlando, Florida Excluded: Abstract not available

Schneider, S., M. Prince-Paul, et al. (2004), Virtual reality as a distraction intervention for women receiving chemotherapy. Onc Nurs Society, Excluded: Abstract not available


Shah, J., D. Buckley, et al. (2003), "Depth cue reliance in surgeons and medical students" Surgical endoscopy 17(9): 1472-4. Excluded: Not related to education of health professionals


Shin, H., G. Stamm, et al. (2000), "[Basic principles of data acquisition and data processing for construction of high quality virtual models], " Der Radiologe 40(3): 304-12. Excluded: Not related to education of health professionals


Sielhorst, T., T. Blum, et al., "Synchronizing 3D movements for quantitative comparison and simultaneous visualization of actions "Excluded: Abstract not available


Sivak, M., C. Mavroidis, et al. (2009), "Design of a low cost multiple user virtual environment for rehabilitation (MUVER) of patients with stroke.” Studies in health technology and informatics 142: 319-24, Excluded: Not related to education of health professionals


Sohmura, T., N. Kusumoto, et al. (2009), "CAD/CAM fabrication and clinical application of surgical template and bone model in oral implant surgery" *Clinical oral implants research* **20**(1): 87-93, Excluded: Not related to education of health professionals


Sugavanam, S. and V. Devarajan (2003), "Simulation of a preperitoneal mesh in laparoscopic herniorrhaphy \textit{Studies in health technology and informatics} \textbf{94}: 343-5, Excluded: Not related to education of health professionals


Toso, F., C. Zuiani, et al. (2005), "Usefulness of computed tomography in pre-surgical evaluation of maxillo-facial pathology with rapid prototyping and surgical pre-planning by virtual reality" La Radiologia medica 110(5-6): 665-75, Excluded: Not related to education of health professionals


Wickbom, G. (1995). "[New technology in medicine and education, It is possible to simulate everything in "virtual reality"][". Lakartidningen 92(32-33): 2897-9, Excluded: Published before 1999


Wolf, I., M. Hastenteufel, et al. (2003), Clinical application of new 3D and 4D visualization and quantification tools for cardiac diagnosis and therapy. Elsevier, Excluded: Abstract not available


Wu, H. and J. Han (2008), "The diagnostic value of multislices spiral CT in ureter transitional cell carcinoma" Yixue Yingxiangxue Zazhi Journal of Medical Imaging 18(12): 1436-1438. Excluded: Abstract not available


Excluded studies - based on full text (Second stage of the review)

Inappropriate publication

Inappropriate population


Inappropriate population


Inappropriate study design.


Appendix C

The appendix C shows the summaries from individual study appraisals


Purpose of the study:
- Firstly, to define validity and reliability of the virtual reality simulator (Vascular Interventional Surgical Trainer) for assessment of endovascular skills
- Secondly to assess effectiveness of the stimulator as a training tool

Participants:
- Twenty surgeons with extensive experience in open vascular surgical procedures (>100 cases)
- Divided into two groups, based upon their experience in endovascular procedures.
- Primary operator - 8 surgeons had performed >50 endovascular procedures
- Remaining 12 surgeons had limited experience in endovascular techniques (<10 procedures).

Intervention:

- Primary operator - experience surgeons (n=8)
- Surgeons with limited experience (n=12)
- Familiarised to the VIST simulator and the task
- Two repetitions - renal artery angioplasty and stent procedure
- Further four sessions on the simulator
- Comparison of performance
- Vascular Interventional Surgical Trainer, Mentice Corporation, Gothenburg, Sweden (VIST simulator)
- Stimulator comprises an interface device, a high-performance desktop computer and two display screens.

**Comparison group:**
- 12 surgeons had limited experience in endovascular techniques (<10 procedures).

**Outcome:**
- The outcome measured for both experts and nonexperts
- Total time taken
- Total amount of contrast fluid used
- Fluoroscopy time
- Benchmark criteria to be achieved (the score for each parameter by experience in endovascular procedures).
- Comparison of performance on the first two sessions was assessed to find out whether the simulated task is construct valid (this substantiate the use of the simulator as a tool to assess endovascular technical skill).
- Learning curve for the novice (clarifies whether repeated practice improves performance toward that of the experienced group).

**Results:**

<table>
<thead>
<tr>
<th>Session used</th>
<th>Group</th>
<th>Total time taken</th>
<th>Amount of contrast fluid used</th>
<th>Fluoroscopy Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1</td>
<td>Experienced</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inexperienced</td>
<td>No difference</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>p value</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session 2</td>
<td>Experienced</td>
<td>571.5s</td>
<td>19.1ml</td>
<td>273s</td>
</tr>
<tr>
<td></td>
<td>Inexperienced</td>
<td>900.0s</td>
<td>42.9ml</td>
<td>441s</td>
</tr>
<tr>
<td></td>
<td><strong>p value</strong></td>
<td>0.039</td>
<td>0.047</td>
<td>0.305</td>
</tr>
<tr>
<td>Over the 6 Sessions</td>
<td>Experienced</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inexperienced</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>p value</strong></td>
<td>0.007</td>
<td>0.021</td>
<td>0.187</td>
</tr>
<tr>
<td>End of the training programme</td>
<td>Experienced</td>
<td>571.5s</td>
<td>19.1ml</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inexperienced</td>
<td>456s</td>
<td>19.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>p value</strong></td>
<td>0.491</td>
<td>0.755</td>
<td></td>
</tr>
</tbody>
</table>
Quality assessment:
- Apart from level of experience there were no other differences between the groups at baseline mentioned, which might, in part, account for any differences in outcome. For example the demographic differences (surgeons with laparoscopic skills would have a shorter learning curve for the acquisition of endovascular skills) between groups were unfortunately not collected, and would have provided interesting comparisons.
- Results are clearly presented and answering the research question. P value is reported where as confident intervals are unavailable.
- The learning curve for the inexperienced group plateaued at the third session. This may be because of the surgeons in the experienced group were not truly experts, or that the tasks on the simulator are too easy for them.
- It would have been desirable to assess the performance of the experienced group on a total of six sessions, to confirm the earlier plateau of their learning curve. However, this was not possible in this study due to timing constraints.

Overall conclusion:
- Conclusions drawn are supported by the study results.
- Surgeons with minimal endovascular experience can improve their time taken and contrast usage during short-phase training on a VR endovascular task. However, it is necessary to corroborate these findings by assessing transfer of these skills to real procedures.
- This study has shown a vascular interventional virtual reality simulator to be a valid tool for both assessment and training in endovascular skills.
- Ethics approval was not necessary for this study, though all surgeons provided informed consent prior to commencement of the trial.
- Question also arises whether the groups were well balanced and whether there were any other differences between the groups at entry to the trial.

**Purpose of the study:**

This article reviews the tools currently available for training and assessment in laparoscopic surgery. Medline searches were performed to identify articles.

**Overall conclusion:**

- Current training involves the use of box trainers with either innate models or animal tissues; it lacks objective assessment of skill acquisition.

- Virtual reality simulators have the ability to teach laparoscopic psychomotor skills, and objective assessment is now possible using dexterity-based and video analysis systems.

Purpose of the study:
- To evaluate a VR ectopic module (LAPSIM), in terms of its validity as a training and assessment tool for gynaecological surgeons
- Setting - Departments of surgery and gynaecology in central London teaching hospitals.

Participants:
- 30 gynaecological surgeons (>50 gynaecological laparoscopic procedures)
- Subdivided into 3 groups based on their level of experience - Novice (<10), intermediate (20–50) and experienced (>100)

Intervention:
- VR simulator (VR ectopic module - LAPSIM; Surgical Science, Gothenburg, Sweden)
- The difficulty of the module can be altered from level 1 (easy) to level 7 (difficult) by setting the size of the pregnancy, the initial bleeding rate (millilitre/second) and the bleeding rate millilitre/second) when a cut is made in tissue.
Comparison group:
- Novice (<10 cases)

Outcome:
- Total time taken to complete each task (seconds)
- Path length of each hand (metre)
- Angular path length of each hand (degrees).
- The software also records total blood loss (millilitre), residual bleeding rate (millilitre/second), ovarian diathermy damage (seconds) and amount of unremoved dissected tissue (if any).

Results:

<table>
<thead>
<tr>
<th>Assessment of performance</th>
<th>Group</th>
<th>Total time taken(s)</th>
<th>Total instrument path length (m)</th>
<th>Total blood loss(ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrument navigation task</td>
<td>Novice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Experienced</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>p value</strong></td>
<td>0.002</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>Session 1</td>
<td>Novice</td>
<td>758.73</td>
<td>18.9</td>
<td>520.9</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>477.385</td>
<td>13.3</td>
<td>328.7</td>
</tr>
<tr>
<td></td>
<td>Experienced</td>
<td>327.71</td>
<td>8.4</td>
<td>142.1</td>
</tr>
<tr>
<td></td>
<td><strong>p value</strong></td>
<td>0.038</td>
<td>0.051</td>
<td>0.060</td>
</tr>
<tr>
<td>Session 2</td>
<td>Novice</td>
<td>551.1</td>
<td>17.8</td>
<td>304.2</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>401.2</td>
<td>8.3</td>
<td>187.4</td>
</tr>
<tr>
<td></td>
<td>Experienced</td>
<td>249.2</td>
<td>6.8</td>
<td>123.3</td>
</tr>
<tr>
<td></td>
<td><strong>p value</strong></td>
<td>0.001</td>
<td>0.023</td>
<td>0.031</td>
</tr>
<tr>
<td>Learning curve plateaus</td>
<td>Novice</td>
<td>ninth session (244.5, P =0.057)</td>
<td>fourth session (median 11.5, P =0.057)</td>
<td>fourth session (183.5, P = 0.118)</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>seventh session (178.0, P = 0.825)</td>
<td>sixth session, (median 7.4, P = 0.099)</td>
<td>seventh session (104.6l, P = 0.789)</td>
</tr>
<tr>
<td></td>
<td>Experienced</td>
<td>second session (249.2 P = 0.084)</td>
<td>second session, (median 6.8, P = 0.258)</td>
<td>third session (median 101.5, P = 0.162)</td>
</tr>
</tbody>
</table>
Quality assessment:

- A total of 23 of the 30 subjects managed to complete all ten sessions on the simulator, and the remainder citing timetabling constraints for their inability to fulfil the demands of the study.
- Limitations to the size of the sample population due to timetabling constraints have led to groups, which were smaller than originally aimed for, and may have affected the statistical analyses.
- There may be other differences between the groups at baseline (e.g., age, previous training on VR stimulator) which might, in part, account for any differences in outcome. Inclusion of confounding variables would have provided interesting comparison.
- No effort was made to achieve blinding and we think it matters in this study as it may introduce observer bias.
- We are not very clear whether outcome measurements have been determined in the same way between study groups and this may lead to measurement or misclassification bias.
- P value is reported whereas confidence intervals are unavailable.
- Ethical issues have been addressed as the aims of the study were explained to all the subjects, and informed consent was obtained prior to participation in the trial.

Overall conclusion:

- The intention of this study was to establish if a VR ectopic module could be a valid teaching tool for trainee gynaecological surgeons. The results have successfully supported this hypothesis.
- The module displays construct validity as the experts performed significantly better than the intermediates and in turn the novices.
- It has thus been proven that training on this simulator can improve laparoscopic procedural skills, but only when measured using the simulator. However, it is necessary to corroborate these findings by assessing transfer of these skills to real procedures.

Purpose of the study:
- Focuses on the importance of haptic feedback
- Find out the additive value of virtual reality simulation in laparoscopic suturing training versus traditional box trainers

Participants:
- Surgical and gynaecology residents with some laparoscopic experience (N = 45)
- Laparoscopic suturing experience was an exclusion criterion for the participation in the arm involving the assessment of the suturing and knot tying skills.

Intervention:
- Incremental value of the test being compared to other routine tests.
- Two types of laparoscopy simulators: two VR simulators and traditional box trainers.
- Both SimSurgery VR simulator systems were without haptics and ran the SimPort software (SEP 1.04.3 SimSurgery, Oslo, Norway).
- Their hardware platform differed: one VR simulator system incorporated the SimPack platform (SimSurgery, Oslo, Norway), while the other system made use of two Xitact HTP instrument ports (Xitact/Mentice SA, Morges, Switzerland)
- The study has two arms:

<table>
<thead>
<tr>
<th>N = 45</th>
<th>Demonstration surgeon’s knot</th>
<th>First part of questionnaire</th>
</tr>
</thead>
</table>

All participants were randomly and blinded divided into two equally sized groups

<table>
<thead>
<tr>
<th>Group A</th>
<th>Group B</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=23</td>
<td>N=22</td>
</tr>
<tr>
<td>Training on laparoscopic suturing on BOX TRAINER - suturing assessment (control group)</td>
<td>Training on laparoscopic suturing on VR SIMULATOR</td>
</tr>
<tr>
<td>Training on laparoscopic suturing on VR SIMULATOR</td>
<td>Training on laparoscopic suturing on BOX TRAINER</td>
</tr>
<tr>
<td>Suturing assessment</td>
<td>Suturing assessment</td>
</tr>
</tbody>
</table>

Remainder of questionnaire
Comparison group:
- Control group was gathered from the data of the assessment of group A (N = 10), after the initial half hour training session on the box trainer

Outcome:
- Performance of suturing skills
  - ‘quality (strength) of knot’
- Questionnaire opinion on the simulators used in the study and their role in laparoscopic suturing training.
  - demographics and prior laparoscopic and simulator experience of the trainees
  - realism and haptic feedback of both simulators
  - preferences of the trainees regarding laparoscopic suturing training.

Results:
- Results clearly presented and answer the research question.

Table 1 - Scoring of the final laparoscopic knot

<table>
<thead>
<tr>
<th></th>
<th>Mean Gr.A</th>
<th>Mean Gr.B</th>
<th>Mean Control</th>
<th>p value A vs control</th>
<th>p value B vs control</th>
<th>p value A vs B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positioning of needle in needle holder</td>
<td>3.80</td>
<td>3.30</td>
<td>3.30</td>
<td>0.145</td>
<td>1.000</td>
<td>0.105</td>
</tr>
<tr>
<td>Running needle through suturing pad</td>
<td>3.90</td>
<td>3.60</td>
<td>3.60</td>
<td>0.232</td>
<td>1.000</td>
<td>0.232</td>
</tr>
<tr>
<td>Taking proper bites of the suturing pad, during suturing</td>
<td>4.10</td>
<td>3.60</td>
<td>3.90</td>
<td>0.441</td>
<td>0.232</td>
<td>0.054</td>
</tr>
<tr>
<td>Throwing thread around needle holder</td>
<td>3.50</td>
<td>3.10</td>
<td>3.20</td>
<td>0.552</td>
<td>0.836</td>
<td>0.400</td>
</tr>
<tr>
<td>Pulling tight of the thread</td>
<td>3.70</td>
<td>3.20</td>
<td>3.70</td>
<td>1.000</td>
<td>0.216</td>
<td>0.216</td>
</tr>
<tr>
<td>Tying a correct ‘surgical knot’</td>
<td>3.50</td>
<td>3.50</td>
<td>3.20</td>
<td>0.512</td>
<td>0.458</td>
<td>1.000</td>
</tr>
<tr>
<td>Quality (strength) of knot (test by pulling on knot)</td>
<td>3.90</td>
<td>3.80</td>
<td>3.60</td>
<td>0.563</td>
<td>0.641</td>
<td>0.773</td>
</tr>
<tr>
<td>Global evaluation of performance</td>
<td>3.90</td>
<td>3.50</td>
<td>3.70</td>
<td>0.574</td>
<td>0.526</td>
<td>0.276</td>
</tr>
<tr>
<td>Summation of scores</td>
<td>30.80</td>
<td>27.60</td>
<td>28.20</td>
<td>0.298</td>
<td>0.772</td>
<td>0.160</td>
</tr>
</tbody>
</table>
Table -2 Opinions on laparoscopic suturing simulators

<table>
<thead>
<tr>
<th>Mean</th>
<th>VR simulator SimSurgery</th>
<th>Traditional box trainer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global impression</td>
<td>3.00</td>
<td>3.95</td>
</tr>
<tr>
<td>Movement of the instruments</td>
<td>2.83</td>
<td>4.37</td>
</tr>
<tr>
<td>Realism of needle and thread</td>
<td>2.75</td>
<td>4.53</td>
</tr>
<tr>
<td>Tying of the knots</td>
<td>2.75</td>
<td>4.53</td>
</tr>
<tr>
<td>Pulling tight of the suturing thread</td>
<td>2.68</td>
<td>4.34</td>
</tr>
<tr>
<td>Movement of the suturing thread</td>
<td>2.71</td>
<td>4.34</td>
</tr>
<tr>
<td>Haptic sensations of the tissue</td>
<td>1.98</td>
<td>3.83</td>
</tr>
<tr>
<td>Resistance of needle and thread</td>
<td>1.93</td>
<td>4.17</td>
</tr>
</tbody>
</table>

Table - 3 Preference of order to practice laparoscopic suturing skills

<table>
<thead>
<tr>
<th>Preference of order to practice laparoscopic suturing skills</th>
<th>Simsurgery first</th>
<th>Box first</th>
<th>Both simulator systems for suturing training</th>
<th>Traditional box trainer alone</th>
</tr>
</thead>
<tbody>
<tr>
<td>53.3%</td>
<td>37.8%</td>
<td>53.3%</td>
<td>46.7%</td>
<td></td>
</tr>
</tbody>
</table>

Quality assessment:
- We are not very clear whether participants recruited to the intervention and control groups in a way that minimised bias and confounders. There may be some differences between the groups at entry to the trial. Inclusion of confounding variables would have provided interesting comparison.
- However, all participants were randomly and blinded divided into two equally sized groups.
- Randomisation stated but method not described
- Participants of both group A (N = 10) and B (N = 10) were equally divided between two objective expert observers, to avoid inter-examiner differences.
- P value is reported where as confident intervals are unavailable.
**Overall conclusion:**

- Laparoscopic suturing, haptic feedback is considered a necessity.
- Participants preferred the box trainer for laparoscopic suturing training until these are also capable to provide sufficient haptic feedback.
- VR simulation does not have a significant additional value in laparoscopic suturing training, over traditional box trainers.
- The conclusion drawn supported by the study results.
- Future development in VR simulation should focus on basic skills and component tasks of procedural training in laparoscopic surgery, rather than laparoscopic suturing.

Purpose:
- To describe VSOne as a multi-media based medical e-learning system
- To evaluate effectiveness of VSOne as a training method

Overview of the paper:
- Chapters of this review have given a system overview with a detailed description of the hardware and software use in VSOne.
- Special emphasis has been given to modelling with the authoring software KisMo, which enables to create patient-specific simulation models.
- Discussed development of several Basic Task Training (BTT) modules to improve dexterity and complex Surgical Procedure Task (SPT) modules for full training of laparoscopic procedures.
- Discussion on the benefits of VR based training systems
- VSOne has been evaluated in a comparative randomised study with classic training methods.

Evaluation study for VSOne

Participants:
- 24 medical students without prior experience and 8 laparoscopic surgeons
- Test persons were randomly assigned to 2 groups with the same group structure

Intervention:

<table>
<thead>
<tr>
<th>VEST group (12 novices, four experts)</th>
<th>CVT group (Conventional Video Training)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual training environment</td>
<td>Mechanical model of the BTT-Tubes and the BTT-Blocks</td>
</tr>
<tr>
<td>Groups switched the devices on the fifth day</td>
<td></td>
</tr>
</tbody>
</table>

Control group:
- CVT group (Conventional Video Training)
Outcome:
- Testing the skills transferability
- Training effect (number of errors, time to complete task)

Results:
- Training effect with VSOne and CVT is comparable; both groups showed typical learning curves.
- Skills acquired with VSOne can be transferred to the real-world environment of a conventional video trainer.
- Laparoscopically experienced surgeons have a notably better performance than novices for both training environments.

Quality assessment:
- Participants recruited to the intervention and control groups in a way that minimised bias.
- Randomisation stated but method not described.
- The key differences between the groups at baseline which might, in part, account for any differences in outcome. Inclusion of confounding variables would have provided interesting comparison.
- We are not very clear whether the participants in all groups followed up and data collected in the same way, reviewed at the same time intervals, in the same way
- Blinding of outcome measurements has not indicated.
- P value and confident intervals are unavailable.

Overall conclusion:
- VSOne guarantees realistic training possibilities for various laparoscopic procedures.
- Evaluation of VSOne has shown the typical learning curves for inexperienced trainees and the transferability of the learning effect to the real-world situation.
- For basic skills training VSOne is suited as well as conventional training.
- For complex surgical procedure training, VR based training is superior due to the outstanding benefits:
  - 24 hour availability
  - reduced training costs
  - variability, repeatability and standardization
  - objective assessment of the laparoscopic training
- Promising trend towards VR based training systems in education and training for laparoscopic surgery.

**Purpose of the study:**
- To find out the effectiveness of the GI Mentor simulator as a training tool

**Participants:**
- 45 first-year Gastroenterology fellows (< 10 colonoscopy cases)

**Intervention:**

Gastroenterology fellows (< 10 colonoscopy cases)

Questionnaire (demographics) and general lectures on colonoscopy

Randomized

Simulator training (group A) n=22

No simulator training (group B) n=23

Five 2-hour sessions on the GI Mentor simulator (10 different cases)

Log form and the questionnaire (impression about the usefulness of the training sessions)

Supervised, actual colonoscopy - (same time both groups)

Evaluation form - technical and cognitive success, and patient comfort level

**Comparison group:**
- No simulator training (group B) n=23
Outcome:

- Primary outcome - comparison of measurement of competency in colonoscopies
  a) Objective competency is the procedure number for the fellow, ability to reach the transverse colon and the cecum without assistance, and the ability to correctly recognize and identify abnormalities
  b) Subjective competency is on a 5-point scale; 1 (totally unskilled) to 5 (competent and expedient).
- Secondary outcome - comparison of patient-discomfort level

Results:

TABLE 1- Longitudinal skill development on GI Mentor simulator

<table>
<thead>
<tr>
<th>Variable</th>
<th>Case 1</th>
<th>Case 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hour 1</td>
<td>Hour 10</td>
</tr>
<tr>
<td>Total procedure time, s</td>
<td>693</td>
<td>301</td>
</tr>
<tr>
<td>Time to cecum, s</td>
<td>239</td>
<td>123</td>
</tr>
<tr>
<td>% of mucosal surface examined</td>
<td>&lt;.001</td>
<td>86.3</td>
</tr>
<tr>
<td>No. episodes of excessive pressure</td>
<td>86.3</td>
<td>82.7</td>
</tr>
<tr>
<td>ES</td>
<td>0.31</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>61.9</td>
<td>85.8</td>
</tr>
<tr>
<td></td>
<td>.004</td>
<td>.004</td>
</tr>
</tbody>
</table>
### TABLE 2 - Comparison between simulator and no-simulator group in objective competence

<table>
<thead>
<tr>
<th>Group</th>
<th>Session 1</th>
<th>S-2</th>
<th>S-3</th>
<th>S-4</th>
<th>S-5</th>
<th>S-6</th>
<th>S-7</th>
<th>S-8</th>
<th>S-9</th>
<th>S-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean in Simulator (n = 23)</td>
<td>50.4</td>
<td>64.5</td>
<td>74.0</td>
<td>76.7</td>
<td>76.8</td>
<td>77.8</td>
<td>80.8</td>
<td>89.5</td>
<td>87.8</td>
<td>92.7</td>
</tr>
<tr>
<td>Mean in no-simulator (n = 22)</td>
<td>40.9</td>
<td>52.0</td>
<td>62.0</td>
<td>64.4</td>
<td>70.2</td>
<td>77.6</td>
<td>80.5</td>
<td>83.7</td>
<td>85.2</td>
<td>90.9</td>
</tr>
<tr>
<td>P value based on t test</td>
<td>.06</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>.03</td>
<td>.91</td>
<td>.89</td>
<td>.01</td>
<td>.02</td>
<td>.04</td>
</tr>
</tbody>
</table>

### TABLE 3 - Comparison between simulator and no-simulator group in subjective competence

<table>
<thead>
<tr>
<th>Group</th>
<th>Session 1</th>
<th>S-2</th>
<th>S-3</th>
<th>S-4</th>
<th>S-5</th>
<th>S-6</th>
<th>S-7</th>
<th>S-8</th>
<th>S-9</th>
<th>S-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean in Simulator (n = 23)</td>
<td>47.6</td>
<td>68.6</td>
<td>76.3</td>
<td>78.0</td>
<td>81.3</td>
<td>82.0</td>
<td>86.1</td>
<td>88.8</td>
<td>88.9</td>
<td>90.8</td>
</tr>
<tr>
<td>Mean in no-simulator (n = 22)</td>
<td>36.6</td>
<td>57.4</td>
<td>68.4</td>
<td>75.4</td>
<td>79.4</td>
<td>82.3</td>
<td>84.1</td>
<td>86.4</td>
<td>86.8</td>
<td>90.5</td>
</tr>
<tr>
<td>P value based on t test</td>
<td>.08</td>
<td>.004</td>
<td>.005</td>
<td>.32</td>
<td>.28</td>
<td>.88</td>
<td>.32</td>
<td>.11</td>
<td>.32</td>
<td>.82</td>
</tr>
<tr>
<td>Group</td>
<td>S- 10</td>
<td>Session 1</td>
<td>S- 2</td>
<td>S- 3</td>
<td>S- 4</td>
<td>S- 5</td>
<td>S- 6</td>
<td>S- 7</td>
<td>S- 8</td>
<td>S- 9</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>-----------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Mean in Simulator 8.9 (n = 22)</td>
<td>25.7</td>
<td>23.2</td>
<td>16.7</td>
<td>16.0</td>
<td>16.7</td>
<td>13.4</td>
<td>11.9</td>
<td>10.5</td>
<td>10.7</td>
<td></td>
</tr>
<tr>
<td>Mean in no-simulator 9.2 (n = 23)</td>
<td>31.4</td>
<td>19.1</td>
<td>19.5</td>
<td>18.2</td>
<td>16.5</td>
<td>13.9</td>
<td>11.3</td>
<td>10.4</td>
<td>11.8</td>
<td></td>
</tr>
<tr>
<td>P value based on t test</td>
<td>.42</td>
<td>.14</td>
<td>.22</td>
<td>.39</td>
<td>.94</td>
<td>.85</td>
<td>.74</td>
<td>.99</td>
<td>.55</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 4 - Comparison between simulator and no-simulator group in discomfort

Quality assessment:
- Selection bias been minimised:
- exclusion criteria for fellows given
- randomisation stated and method described (random-number table)
- Blinding of outcome measurements has indicated. Fellows were given code numbers to identify them on all study forms. Also proctors filling out the individual evaluation forms remained blinded as to whether the particular fellows did or did not receive prior simulator training.
- Outcome measurements were clearly described and determine in the same way between interventions to minimized measurement or misclassification bias.
- P value is reported where as confident intervals are unavailable.
- Results clearly presented and answer the research question.

Overall conclusion:
- Study demonstrated the benefit of simulator training
- Fellows who underwent GI Mentor training significantly shorten the colonoscopy learning curve and performed significantly better during the early phase of real colonoscopy training

**Purpose of the study:**
- To find out the effectiveness of GI Mentor II (virtual reality endoscopic trainer) as a valid tool for the assessment of endoscopic skills

**Participants:**
- In the first part of the study, 28 subjects were divided into three groups on the basis of their experience with gastrointestinal (GI) endoscopy.
  - Group 1 - experienced surgeons performed > 200 endoscopic procedures, (n = 8)
  - Group 2 - residents performed < 50 endoscopic procedures, (n = 10)
  - Group 3 - medical students who never performed GI endoscopy, (n = 10)
- In the second part of the study, 20 subjects (all were novices) who had never performed GI endoscopy.

**Intervention:**
- GI Mentor II with modified Pentax ECS-3840F endoscope
- The simulator software offers three different tasks, each on three levels
  - Task 1 (Endobasket) - navigation through a virtual bowel, picking and placing balls
  - Task 2 (Endobubble) - navigation through the virtual colon and piercing 20 balloons
  - Task 3 (Virtual Endoscopy) - identification of different pathologies among cases
Comparison group:
- Novices who had never performed GI endoscopy

Outcome:
- Assessment of the learning curve on the simulator was based on the following three parameters: time used, number of punctured balloons, and number of wall collisions.
- Assessment of endoscopic skills during the colonoscopy (2nd part of the study) was based on parameters measured by the computer system: time, percent of mucosa surface examined, efficiency of screening, time with clear view, excessive local pressure, pain, time with pain, loop formation, and total time with loop.
Results:
Table 1 - Learning curves

<table>
<thead>
<tr>
<th></th>
<th>Learning curves plateau</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experienced surgeons</td>
<td>after the 2nd repetition</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>Residents</td>
<td>after the 5th repetition</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>Medical students</td>
<td>after the 7th repetition</td>
<td>&lt; 0.05</td>
</tr>
</tbody>
</table>

Table 2 - Comparison of performance scores for the group that received psychomotor training vs the control group.

<table>
<thead>
<tr>
<th></th>
<th>Psychomotor training</th>
<th>P value</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before</td>
<td>After</td>
<td>Before</td>
</tr>
<tr>
<td>Time (min)</td>
<td>4.5</td>
<td>2.8</td>
<td>5</td>
</tr>
<tr>
<td>% mucosa surface examined</td>
<td>80</td>
<td>85</td>
<td>80</td>
</tr>
<tr>
<td>Efficiency of screening %</td>
<td>85</td>
<td>90</td>
<td>85</td>
</tr>
<tr>
<td>Time with clear view %</td>
<td>92</td>
<td>98</td>
<td>95</td>
</tr>
<tr>
<td>Time with pain %</td>
<td>19</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>Time with loop (min)</td>
<td>1.4</td>
<td>.4</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Quality assessment:
- Participants were appropriately allocated to intervention and control groups.
  Randomization schedule was generated using closed envelopes.
- Researchers did not produce demographic information separately for the intervention and control groups. This may result in dissimilar or unbalanced groups and may have introduced selection bias into the study.
- Outcome measurements were clearly described and determine in the same way between interventions
- Blinding of outcome measurements has not indicated. This becomes more crucial as the measures are more subjective and hence more open to observer bias.
- The results clearly presented and answer the research question.
- P value is reported where as confident intervals are unavailable

Overall conclusion:
- Conclusion drawn supported by the study results.
- The learning rate on the simulator was proportional to the endoscopic experience. (System assesses performance parameters with clinical relevance).
- Residents with limited endoscopic experience as well as beginners will benefit from training in a virtual environment, thereby improving their psychomotor skills.
- Experience group would not benefit from training on the simulator.
This evidence can be considered when designing GI endoscopy training programs in the future.

Study indicates that the GI Mentor is a valid tool for the evaluation of psychomotor skills in gastrointestinal endoscopy.

Furthermore, it shows that psychomotor training has a significant impact on the performance of a simulated colonoscopy.

It has thus been proven that training on this simulator can improve endoscopic procedural skills, but only when measured using the simulator. However, it is necessary to corroborate these findings by assessing transfer of these skills to real procedures.

**Purpose of the study:**
- To assess the discriminative validity of the Minimally Invasive Surgical Trainer in Virtual Reality (MIST-VR) using criteria levels based on expert performance

**Participants:**
- 100 medical students with no laparoscopic operative experience
- 12 experienced laparoscopic surgeons who had performed >50 laparoscopic operations
- 12 less experienced surgeons who had performed more than one but <10 laparoscopic procedures
- 12 university students (novices) who had no medical background

**Intervention:**
- MIST-VR system used in this study was based on a 200-MHz Pentium personal computer (PC) running Windows 95 with 32-Mb RAM and a Matrox Mystique 4-MB video card

**Comparison group:**
- university students (novices) who had no medical background
Outcome:
- Measures of the participants’ performance:
  - time to complete all six MIST-VR tasks
  - number of errors,
  - economy of movement
  - economy of diathermy

Results:

Table 1: Comparison to the control, less experienced, and experienced groups on trial three

<table>
<thead>
<tr>
<th></th>
<th>Novices</th>
<th>Medical students</th>
<th>Less experienced</th>
<th>Experienced</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean number of seconds</td>
<td>900</td>
<td>900</td>
<td>1000</td>
<td>800</td>
<td>p = 0.0058 (medical students / experienced)</td>
</tr>
<tr>
<td>Mean error scores</td>
<td>17</td>
<td>17</td>
<td>19</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Mean economy of movement Scores-right</td>
<td>20</td>
<td>20</td>
<td>24</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Mean economy of movement Scores-left</td>
<td>22</td>
<td>21</td>
<td>22</td>
<td>19</td>
<td>p = 0.02 (medical students / experienced)</td>
</tr>
<tr>
<td>Mean economy of diathermy scores</td>
<td>3.5</td>
<td>2.8</td>
<td>3</td>
<td>2.6</td>
<td>p = 0.003 (medical students /less experienced)</td>
</tr>
</tbody>
</table>

Table 2: Three trials completed by medical students

<table>
<thead>
<tr>
<th></th>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Trial 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean number of seconds</td>
<td>1500</td>
<td>1100</td>
<td>900</td>
</tr>
<tr>
<td>Mean error scores</td>
<td>25</td>
<td>17</td>
<td>16</td>
</tr>
</tbody>
</table>
Quality assessment:
- Researchers produce key demographic information separately for the intervention and control groups.
- No effort was made to achieve randomisation when allocating intervention and control groups.
- Outcome measurements were clearly described and determine in the same way between the groups.
- Blinding of outcome measurements has not indicated. This may introduce observer bias.
- Results clearly presented and answer the research question.
- P value is reported where as confident intervals are unavailable.

Overall conclusion:
- The MIST-VR can measure psychomotor ability, as well as the variability in performance between subjects with similar experience.
- This discrimination among levels of technical ability may be useful in evaluating and training laparoscopic surgeons.
- The MIST-VR may be useful in identifying that subset of novices.

**Purpose of the study:**
- To investigate the value of virtual reality (VR) airway simulator for training

**Participants:**
- Nineteen anaesthesiologists, 15 residents, and 4 attending physicians, participated in the study.
- 4 experienced attending anaesthesiologists served as the experts group having a minimum of 50 fiberoptic intubation (FOI)
- 11 residents for the novice training group who did not have any clinical experience in FOI before the start of this study.
- 4 residents who were unable to use the VR airway simulator throughout the one-week training period served as the novice non-training control group for FOI in the cadaver.

**Intervention:**
- Device - AccuTouch Bronchoscopy Simulator (Immersion Medical, Gaithersburg, MD)
- It is a VR partial-task trainer consisting of a proxy flexible bronchoscope, a robotic interface, a computer, a monitor, and comprehensive simulation software capable of delivering different realistic bronchoscopy scenarios.

---

**Diagram:**

```
Residents group
  Pre-training assessment of performance on VR airway simulator - adult FOI scenario
  Post-training assessment of performance - adult FOI scenario
  Subjective assessment questionnaire
  Assessment of FOI performance in a cadaver

Experienced anaesthesiologists
  FOI performance with the same adult VR FOI scenario
```
**Comparison group:**
- Residents (novice training group) did not have any clinical experience in FOI before the start of this study

**Outcome:**
- Time to intubation before and after a 4-day training period using an adult VR FOI scenario
- Time to intubation using a fresh human cadaver two weeks after the training experience

**Results:**

Table 1: Comparison of time to intubation in the adult VR scenario separated by groups

<table>
<thead>
<tr>
<th>Experts</th>
<th>Novices training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial assessment;</td>
<td></td>
</tr>
<tr>
<td>duration of FOI (s)</td>
<td></td>
</tr>
<tr>
<td>136</td>
<td></td>
</tr>
<tr>
<td>79 +/- 18.2</td>
<td>114 +/- 32.3</td>
</tr>
<tr>
<td>(50-108)</td>
<td>(93-)</td>
</tr>
<tr>
<td>Final assessment;</td>
<td></td>
</tr>
<tr>
<td>duration of FOI (s)</td>
<td></td>
</tr>
<tr>
<td>43-101</td>
<td>62-87</td>
</tr>
<tr>
<td>72 +/- 18.0</td>
<td>75 +/- 18.8</td>
</tr>
<tr>
<td>(43-101)</td>
<td>(62-87)</td>
</tr>
<tr>
<td>p value</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td>0.001</td>
<td>p &lt;</td>
</tr>
</tbody>
</table>

Table 2: Comparison of time to intubation in the cadaver separated by groups

<table>
<thead>
<tr>
<th>Experts</th>
<th>Novices training</th>
<th>Novices non-training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of FOI (s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 +/- 4.6</td>
<td>24 +/- 5.3</td>
<td>86 F 38</td>
</tr>
<tr>
<td>(20-27)</td>
<td>(25-147)</td>
<td></td>
</tr>
<tr>
<td>p value</td>
<td>p &gt; 0.05</td>
<td>p &lt; 0.001</td>
</tr>
</tbody>
</table>
Quality assessment:
- This is an Observational study with pre-post training assessments done in University anaesthesiology department setting.
- Only 4 experienced attending anaesthesiologists served as the experts group.
- Although the method of allocation was described, no details have given to show that attempt has made to minimise the selection bias.
- There may be some differences between the groups at entry to the trial. The researchers not produce demographic information.
- Outcome measurements have been determine in the same way between study groups.
- However, blinding of outcome measurements has not indicated.
- Results clearly presented and answer the research question.
- P value is reported where as confident intervals are unavailable.

Overall conclusion:
- The results indicate that residents can be taught basic FOI skills to a degree that compares to physicians who have been trained using traditional methods.
- The fact that residents with training using the VR airway simulator improved their skills in contrast to non-training residents proves that this improvement can be attributed to the VR airway simulator.
- Experts with advanced FOI skills are unlikely to improve their performance significantly using this VR airway simulator.
- Simulator assessment could be used to compare novices’ performance with experts’ performance in the same way that cadaver assessment is used.
- Results indicate that VR airway simulator assessment can be used for assessment of FOI performance.

Purpose of the study:
- To demonstrate correlation between surgical experience and performance on a virtual reality arthroscopy simulator
- To evaluate the consistency of the simulator performance

Participants:
- Medical students and residents
- 5 subjects with no initial experience on shoulder arthroscopy
- 5 subjects with limited experience (on average, 20 shoulder arthroscopic surgeries)

Intervention:
- Procedicus arthroscopy simulator (Mentice Corp, Göteborg, Sweden), consisting of 6 repetitions of the same training module
- The module required the participant to locate and probe a simulated target within the shoulder joint.

Comparison group:
- Another subject group of similar experience from author’s previous publication (historic controls, N = 14)
Outcome:
- Simulator calculates the following 4 parameters:
  - time to completion of the module
  - number of probe collisions with the tissues
  - average velocity of probe movement
  - distance travelled with the tip of the simulated probe compared to an optimal computer-determined distance
- In addition, to evaluate consistency of simulator performance, results were compared to historical controls of equal experience

Results:
TABLE 1 - Performance Scores of the Same Group and Historical Control Group

<table>
<thead>
<tr>
<th></th>
<th>Baseline Historical Scores Group</th>
<th>Current Scores</th>
<th>% Change</th>
<th>P Value</th>
<th>Control Average</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to completion, s</td>
<td>101.4 ± 1</td>
<td>49.3 ± 1</td>
<td>−51</td>
<td>&lt; .001</td>
<td>59.9</td>
<td></td>
</tr>
<tr>
<td>Hook collisions, n</td>
<td>34.7 ± 1</td>
<td>24.6 ± 1</td>
<td>−29</td>
<td>.025</td>
<td>21.2</td>
<td>.4</td>
</tr>
<tr>
<td>Probe velocity, cm/s</td>
<td>0.22 ± 0.001</td>
<td>0.49 ± 0.001</td>
<td>122</td>
<td>&lt; .001</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Actual more than optimal path, n</td>
<td>7.7 ± 0.001</td>
<td>5.2 ± 0.001</td>
<td>−32</td>
<td>.003</td>
<td>4.9</td>
<td>.5</td>
</tr>
</tbody>
</table>

Quality assessment:
- This is a 3-Year Follow-up controlled laboratory Study.
- Study included a relatively small sample of subjects who were retested after gaining additional arthroscopic experience
- 10 subjects (out of 19 participants from author’s original study) were selected for the current study. However, allocation process was not truly random. We are not very sure that 10 subjects retested differed meaningfully from the 9 subjects could not retest.
- Researchers do not produce key demographic information separately for the intervention and control groups.
- We are not very clear whether adequate adjustments been made for residual confounding
- Outcome measurements were clearly described. However, no effort was made to achieve blinding.
• However, researchers do not make clear what had happen to all the participants at the end of 3-Year Follow-up
• Results clearly presented and answer the research question. Conclusion drawn supported by the study results
• P value is reported where as confident intervals are unavailable

Overall conclusion:
• Subjects significantly improved their performance on simulator retesting 3 years after initial evaluation.
• These results further validate the use of surgical simulation as an important tool for the evaluation of surgical skills.
• Simulator yields consistent results across groups with similar surgical experience
• Additionally, it may be possible to establish simulator benchmarks to indicate likely arthroscopic skill.

**Purpose of the study:**
- To validate the role of VR simulation as a tool for surgical skills training

**Participants:**
- Twenty surgeons with limited experience in laparoscopic surgery (median 4.5 (range 0–8) cholecystectomies)

**Intervention:**
- VR training included ten repetitions of all six tasks of the Minimally Invasive Surgical Trainer – Virtual Reality (MIST-VR; MenticeMedical Simulation, Gothenburg, Sweden)
- The six tasks are of progressive complexity and are designed to simulate the techniques used during laparoscopic cholecystectomy.

**Comparison group:**
- Group assigned to no training (n = 10)
Outcome:
- Duration of procedure (min)
- Error score
- Economy of movement score

Results:

<table>
<thead>
<tr>
<th></th>
<th>VR training</th>
<th></th>
<th>Control</th>
<th></th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Operation 1</td>
<td>Operation 2</td>
<td>Operation 1</td>
<td>Operation 2</td>
<td></td>
</tr>
<tr>
<td>Duration of procedure (min)</td>
<td>65</td>
<td>55</td>
<td>56</td>
<td>58</td>
<td>0·021</td>
</tr>
<tr>
<td>Error score</td>
<td>6.4</td>
<td>3</td>
<td>6</td>
<td>5.5</td>
<td>0·003</td>
</tr>
<tr>
<td>Economy of movement score</td>
<td>5.8</td>
<td>3</td>
<td>6</td>
<td>5.9</td>
<td>0·003</td>
</tr>
</tbody>
</table>

Quality assessment:
- Sample size was small; a larger trial is required to confirm the present findings.
- Randomisation claimed and described (performed using sealed envelopes).
- Researchers did not produce demographic information separately for the intervention and control groups.
- The reviewers were blinded to the training status of the trainees and performed the evaluation independently.
- Researchers make clear what had happen to all the participants at the end of the study. Four of the trainees (two in the VR training group and two in the control group) had technical problems (no image on the tape) and were excluded.
- Outcome measurements were clearly described and determine in the same way between two groups.
- The results clearly presented and answer the research question.
- P value is reported where as confident intervals are unavailable.

Overall conclusion:
- Training in a virtual environment can contribute to the development of technical skills relevant to the performance of laparoscopic surgery in vivo.
- The present study providing strong evidence for the role of MIST-VR laparoscopy trainer as a valid tool for training laparoscopic psychomotor skills.
Purpose of the study:
- To validate the Minimally Invasive Surgical Trainer, Virtual Reality (MISTVR) as an objective method for assessing laparoscopic surgical skills

Participants:
- Fourteen surgical residents (13 male, 1 female) with similar limited experience in endoscopic surgery (<10 cholecystectomies).

Intervention:
- Laparoscopic skills in vitro were measured objectively by performing the six tasks on the MIST-VR system (Virtual Presence Ltd., London, England).
- The system is based on a PC and configured with a Pentium 200-MHz processor, 32 MB of RAM, a 1.6-GB hard drive.
- It is linked to a frame containing two laparoscopic instruments and a diathermy pedal.

Comparison group:
- No identifiable comparison group
Outcome:
- During the MIST-VR tasks:
  - errors (number of movements away from the target)
  - non-economy of movement for each hand (actual path length/ideal path length)
  - operation time
- Laparoscopic skill in vivo (laparoscopic cholecystectomy) on anesthetized pigs was assessed according to predefined objective criteria:
  - coordination
  - confidence of movements
  - bleeding
  - gallbladder perforation
  - deep lesions
  \{ total economy score \}
  \{ total error score \}

Results:
Table 1 - Performance scores for the MIST-VR (task 6, third test session) and the animal operation.

<table>
<thead>
<tr>
<th>MIST-VR Economy No.</th>
<th>Economy Errors</th>
<th>R hand Economy</th>
<th>Error</th>
<th>L hand Economy</th>
<th>Animal operation Economy score</th>
<th>Animal operation Economy score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.00</td>
<td>5.46</td>
<td>7.76</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7.75</td>
<td>4.28</td>
<td>6.92</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>16.75</td>
<td>6.47</td>
<td>11.13</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6.00</td>
<td>4.20</td>
<td>5.66</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3.25</td>
<td>3.85</td>
<td>5.16</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7.00</td>
<td>5.03</td>
<td>5.95</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4.75</td>
<td>4.36</td>
<td>6.14</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>7.50</td>
<td>3.56</td>
<td>5.39</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>6.25</td>
<td>13.41</td>
<td>4.28</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.00</td>
<td>4.25</td>
<td>2.63</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>3.50</td>
<td>5.15</td>
<td>3.89</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1.75</td>
<td>3.95</td>
<td>4.46</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1.50</td>
<td>3.76</td>
<td>2.99</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>2.25</td>
<td>2.44</td>
<td>2.77</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task</td>
<td>Error scores</td>
<td>Economy score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------------</td>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>animal procedure and three of the six virtual tasks</td>
<td>animal procedure and non-economy of motion scores for the right hand of the simulator tasks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$p$ value</td>
<td>$p$ value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 1</td>
<td></td>
<td>0.034</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 2</td>
<td>0.012</td>
<td>0.014</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 4</td>
<td>0.049</td>
<td>0.018</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 5</td>
<td></td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 6</td>
<td>0.038</td>
<td>0.022</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Quality assessment:**
- Sample size in this study is small; thus, larger studies are needed to confirm these findings.
- Scoring of the performances in the animal operation was subjective. This can introduced bias to the study. However, authors have taken steps to minimize this by establishing objective and easily assessable scoring criteria.
- Cholecystectomy is considered to be easier to perform in pigs compared with humans. Future studies should therefore correlate simulator performance with performance in the operating theatre.
- $P$ value is reported where as confident intervals are unavailable

**Overall conclusion:**
- The present study demonstrates that in vitro scores for tasks performed on a MISTVR are comparable to performance during operations on living animals.
- Thus, computer model represents a promising tool for the evaluation of junior surgical residents, and can provide them with an objective quantitative assessment of their laparoscopic skills.

Purpose of the review:

- To determine whether virtual reality (VR) training can supplement and/or replace conventional laparoscopic training in surgical trainees with limited or no laparoscopic experience.

Search strategy:

- This review included a methodical search of all the relevant medical, educational and computer literature databases including the grey literature24

Selection criteria:

- Authors included all randomised clinical trials comparing virtual reality training versus other forms of training including video trainer training, no training, or standard laparoscopic training in surgical trainees with little or no prior laparoscopic experience.
- They also included trials comparing different methods of virtual reality training.

Flow chart of articles identified, included and excluded:

References identified through electronic searches of databases (n = 2176) → References excluded (n = 2125) Duplicates (n = 1022) Irrelevant by reading titles and abstracts (n = 1103)

References retrieved for more detailed evaluation (n = 51) → References excluded (n = 21) Compares the ability of VR model and VT model in distinguishing experts and novices (n = 1) Compares local versus telementoring (n = 1) Not VR training model (n = 15) Not RCT (n = 4)

Potentially appropriate references to be included in the systematic review (n = 30)
Outcomes:
Primary outcomes
1. Patient or animal mortality
2. Patient or animal morbidity

Secondary outcomes
1. Conversion to open procedure.
2. Operating time (post training).
3. Time taken to perform the evaluation task on the simulation model (post training).
4. Error score
5. Accuracy.
6. Composite score.
7. Movements (distance and error).
8. Participant satisfaction.

Author’s conclusion:
- The trials included in this review mainly assessed the effect of VR training on the development or improvement of generic skills.

Findings:
- VR training can supplement standard laparoscopic surgical training.
- It is at least as effective as video training in supplementing standard laparoscopic training.
- There was no comparison between different VR models.
- None of the trials assessed VR training as part of a surgical training curriculum.
- VR training without haptic feedback is not realistic.
- Recent advances in VR technology have made it possible to import images into VR software from external sources.
Hassan, I., M. Koller, et al. (2006). "Improvement of surgical skills after a three-day practical course for laparoscopic surgery." Swiss medical weekly : official journal of the Swiss Society of Infectious Diseases, the Swiss Society of Internal Medicine, the Swiss Society of Pneumology 136: 631-636.

**Purpose of the study:**
1. To evaluate if a three day practical course for laparoscopic surgery can improve laparoscopic skills of residents in surgery.
2. To determine which degree of existing laparoscopic experience is required to achieve the most benefit from such a practical course in laparoscopic surgery.

**Participants:**
- Test group 44 surgical residents
  - Divided into two groups:
    1. Advanced participants: 18 residents (> 50 laparoscopic operations)
    2. Novices: 26 residents (< 10 laparoscopic operations)
- Gold standard control group comprised of surgical consultants (n = 6)
- Second control group - medical students in their final year (n = 20) with limited laparoscopic experience
- All participants of the study were without any previous experience with a virtual reality simulator.

**Intervention:**
- The simulator used in this study (LapSim, Surgical Science Ltd., Goteborg/Sweden) creates a virtual laparoscopic system using a computer (Windows XP), a video monitor and laparoscopic interface containing two pistolgrip instruments and a diathermy pedal without haptic feedback.
- The LapSim software contains the basic modules referred to as “clip-and-cut task”, in which the level of complexity and difficulty can be adjusted as previously described.
Comparison group:
- Medical students in their final year (n = 20) with limited laparoscopic experience.
- Gold standard control group comprised of surgical consultants (n = 6)

Outcome:
- Time needed to complete the task (min)
- Error score
- [blood loss (dl)]
- Dropped clips (n)
- Badly placed clips (n)
- Incomplete target areas (n)
- Economy of motion
- Instrument path length (m) and angular path (°)
Results:

Fig. 1 – Mean and SD of the time in Seconds of the clip application on LapSim

Fig. 2 – Mean and SD of the error scores of the clip application on LapSim
Quality assessment:

- Participants recruited to the intervention and control groups in a way that minimised bias and confounders.
- Outcome measurements were clearly described and determined in the same way between interventions.
- Results are clearly presented and answer the research question.
- No effort was made to achieve blinding. We think it matters in this study as it may introduce observer bias.

Overall conclusion:

- A three-day practical course for laparoscopic surgery improved laparoscopic skills of residents.
- However, advanced residents benefit most from the course.

**Purpose of the study:**
- To evaluate the effectiveness of the UroMentor trainer as a training tool
- To evaluate the actual improvement in clinical endourological surgery after simulator training

**Participants:**
- Twenty experienced urologists (total flexible URSs, 21–153) with no previous experience with virtual URS
- A second group comprised five urological residents with no endourological experience

**Intervention:**
- Study used computer based simulator for semi-rigid and flexible ureterorenoscopy (URS), as previously described (UroMentor, Simbionix Ltd, Israel)

Twenty experienced urologists

Training session on the simulator

Undertaken simulated flexible URS for treating a lower calyceal stone

Outcome was correlated with individual experience

Clinical skills compared to residents of a similar level of experience, with no former simulator training, during endoscopic procedures in five comparable cases

Five urological residents with no endourological experience

10 simulator training sessions

Acquired clinical skills compared to those of five urological residents who received no simulator training
Comparison group:
- Urological residents who received no simulator training

Outcome:
- Total operation time
- X-ray exposure
- Guide wire insertion time
- Time of progression from the orifice to the stone
- Stone contact time
- Number of perforations
- Bleeding events laser misfiring
- Scope damage and treatment success

Results:

Table 1 - Overview of performance variables achieved by subgroups with different clinical experience

<table>
<thead>
<tr>
<th>Previous experience (number of previous flexible URSs)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 80</td>
<td>60-80</td>
</tr>
<tr>
<td>Mean operation times</td>
<td>12.3</td>
</tr>
</tbody>
</table>

Table 2 - The reduction in operating time for residents trained on the simulator and untrained

<table>
<thead>
<tr>
<th>Reduction in the operating time after 5 cases</th>
<th>Simulator trained residents</th>
<th>Untrained residents</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9</td>
<td>5</td>
<td>&lt; 0.05</td>
</tr>
</tbody>
</table>

Quality assessment:
- No details have given to show that attempt has made to minimised the selection bias
- Randomisation not mentioned
- Outcome measurements were clearly described and determine in the same way between the groups
- The results clearly presented and answer the research question
- P value is reported where as confident intervals are unavailable
- Blinding is not stated. Supervisors were aware of whether the resident had had simulator training before.
Overall conclusion:
- Individual experience correlates with individual performance on the simulator.
- Simulator training is helpful in improving clinical skills.
- Virtual reality-based training has the potential to become an important tool for clinical education.

**Purpose of the study:**
- To test the ability of the LapSim laparoscopic simulator to assess existing clinical laparoscopic experience
- Thereby further validate its use in surgical assessment and education.

**Participants:**
- 115 surgeons with laparoscopic experience

**Intervention:**
- Laparoscopic skills were objectively measured by performing tasks on the LapSim laparoscopic simulator system
- It is based on a PC, linked to a jig containing two laparoscopic instruments and a diathermy pedal, and movement is translated into a real-time graphic display.
- The system used does not possess haptic feedback.

Comparison groups
- Group 1, novices <50 operations (n=61) and group 2, experts >50 operations (n=54)

**Outcome**
- Time to perform each task
- Path lengths of the instruments and their respective angles
- Tissue damage
Results

Table 1: Timings of the exercises

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Group 1, &lt;50 operations</th>
<th>Group 2, &gt;50 operations</th>
<th>Difference (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Navigation</td>
<td>29.7 (6.36–151.4)</td>
<td>26.1 (13.2–69.3)</td>
<td>12.1</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Grasping</td>
<td>78.6 (20.1–329.1)</td>
<td>63.2 (25.2–146.1)</td>
<td>19.6</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Coordination</td>
<td>56.2 (10.6–282.4)</td>
<td>50.9 (16.9–305.9)</td>
<td>9.4</td>
<td>0.12</td>
</tr>
<tr>
<td>Cutting</td>
<td>119.4 (31.5–463.6)</td>
<td>71.8 (38.8–177.9)</td>
<td>39.9</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Clipping</td>
<td>113.7 (31.0–549.9)</td>
<td>91.3 (34.3–368.9)</td>
<td>19.7</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Total time</td>
<td>424.3 (99.4–1,376.3)</td>
<td>315.3 (168.4–625.4)</td>
<td>25.6</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

Table 2: Results of the cutting exercise

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Group 1, &lt;50 operations</th>
<th>Group 2, &gt;50 operations</th>
<th>Difference (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrument pathways</td>
<td>1.42 (0.59–7.33)</td>
<td>1.18 (0.69–2.52)</td>
<td>17</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Grasping (m)</td>
<td>0.85 (0.34–4.59)</td>
<td>0.76 (0.41–1.76)</td>
<td>9.5</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Coagulation device (m)</td>
<td>352 (104–1,628)</td>
<td>273 (149–672)</td>
<td>22</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Grasping (°)</td>
<td>208 (27–1,021)</td>
<td>204 (107–444)</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>Tissue damage</td>
<td>3 (0–32)</td>
<td>2 (0–13)</td>
<td>33</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Quality assessments

- All participants were stratified into two groups depend only on their experience. There may be other differences between the groups at entry to the trial. We are not very clear whether adequate adjustments been made for residual confounding. E.g. it was not recorded whether participants had had prior exposure to laparoscopic simulators. However, to rule out a possible influence of the chosen grouping researchers carried out additional data analysis, comparing novices with surgeons having > 100 laparoscopic procedures.
- Outcome measurements were clearly described and determine in the same way between interventions.
- No effort was made to achieve blinding. We think it matters in this study as it may introduce observer bias.
- Results clearly presented and answer the research question.
- P value is reported where as confident intervals are unavailable.
- Researchers did not make clear what had happen to all the participants at the end of the study

Overall conclusion

- It can be concluded that laparoscopic skills acquired in the operating room transfer into virtual reality.
- Conclusions drawn are supported by the study results.
- A laparoscopic simulator can serve as an instrument for the assessment of experience in laparoscopic surgery.

**Purpose of the study**
- Proficiency based virtual reality training in laparoscopic salpingectomy compared with standard clinical education
- To assess the effect of virtual reality training on an actual laparoscopic operation

**Participants**
- 24 first and second year registrars specialising in gynaecology and obstetrics
- Participants had no experience of advanced laparoscopy

**Intervention**
- Virtual reality laparoscopy simulator program (LapSim Gyn v 3.0.1; Surgical Science, Gothenburg, Sweden)
- It was run on an IBM T42 computer in a docking station (PentiumM1.8 GHz/512 MB RAM; IBM Armonk, NY, USA) using an interface with a diathermy pedal (Virtual Laparoscopic Interface; Immersion, San Jose, CA, USA).
Comparison groups
- Simulator trained group (n=11) and control group with standard clinical education (n=10)

Outcome
- The main outcome measure was technical performance
- The secondary outcome measure was operation time in minutes

Results

![Table showing impact of virtual reality simulator training on surgical performance and operation time. Values are medians (ranges; interquartile ranges) unless stated otherwise.]

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Simulator trained group (n=11)</th>
<th>Control group (n=10)</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgical performance:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score (points)</td>
<td>33 (25-39; 32-36)</td>
<td>23 (21-28; 22-27)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>% reaching ≥30 points</td>
<td>82</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Operation time:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total time (minutes)</td>
<td>12 (6-24; 10-14)</td>
<td>24 (14-38; 20-29)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Inter-rater agreement 0.79, y-coefficient 0.83 (95% confidence interval 0.68 to 0.98).</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Quality assessments
- Design prospective randomised controlled and blinded trial
- Randomisation claimed and described. However, owing to the nature of the trial trainees were not blind to their allocated group
- Groups were well balanced. Baseline characteristics including demographic information and prior exposure to simple laparoscopy of gynaecology trainees at entry to the trial were reported
- Researchers made clear what had happened to all the participants at the end of the study
- Outcome measurements were clearly described and participants in all groups followed up and data collected in the same way
- Results did answer the research question

Overall conclusion
- It is possible to transfer skills acquired during proficiency based training using a virtual reality simulator to a real operation
**Purpose of the study**
- To demonstrated the VR learning curve for novice surgeons performing endovascular skills tasks
- To investigate transference of skills between VR endovascular skills tasks

**Participants**
- Twenty surgeons with no previous endovascular experience were randomised into two equal groups
- Group A - VR iliac training module and Group B - VR renal training module.

**Intervention**
- Vascular Intervention System Training simulator (VIST, Mentice Corporation, Gothenderg, Sweden)
- The Simulation software uses reconstructions based on real contrast enhanced CT scans and force feedback

---

![Diagram](image.png)

**Basic Surgical Trainees (n=20)**

**Protocol A (n=10)**
- Task 1: Simplified simulation of iliac artery angioplasty (8 repetitions)

**Protocol B (n=10)**
- Task 2: Simplified simulation of renal artery angioplasty (8 repetitions)

**Cross over**
- Assessed on task 2 - renal angioplasty (2 repetitions)
- Assessed on task 1 - iliac angioplasty (2 repetitions)

**Outcome**
- Procedure time (PT), fluoroscopy time (FT), contrast fluid used (CF), placement accuracy (PA), residual stenosis (RS), lesion coverage (LC), stent-vessel ratio (SVR) and maximum stent deployment pressure (MP)
Results

- Over the eight sessions both groups demonstrated statistically significant VR learning curves.
- At crossover, subjects in group B (renal trained) performed to the same level of skill on the simulated iliac task as group A.
- However, those in group A (iliac trained) had a significantly higher fluoroscopy time (median 118 vs 72 secs, p= 0.020) when performing their first simulated renal task than for group B.

Quality assessments

- Randomisation stated and method described.
- The researchers did produce a table listing key demographic information and other level of training.
- Outcome measurements were clearly described and determine in the same way between interventions.
- Blinding of outcome measurements has not indicated. This becomes more crucial as the measures are more subjective and hence more open to observer bias.
- P value is reported where as confident intervals are unavailable.
- Conclusion drawn supported by the study results.

Overall conclusion

- Novice endovascular surgeons can significantly improve their performance of simulated procedures through repeated practice on VR simulators.
- Skills transfer between tasks was demonstrated but complex task training, such as selective arterial cannulation in simulators and possibly in the real world appears to involve a separate skill.

Purpose of the study
- To evaluate AccuTouch Flexible Bronchoscopy Simulator

Participants
- Twenty pediatric residents with no prior experience in bronchoscopy

Comparison groups
- 12 residents in the “Simulator” group, and 8 residents in the “Control” group, randomly assigned.

Intervention
- The Simulator (AccuTouch Flexible Bronchoscopy Simulator) consists of a proxy flexible bronchoscope, a robotic interface device, computer with monitor, and simulation software.

Outcome
- Intubations were analyzed for: time to visualization of the carina, and number and time that the bronchoscope tip hit the mucosa.
Results

### Table 2: Comparison of Simulator and Control Intubation Variables

<table>
<thead>
<tr>
<th></th>
<th>1st Simulator Mean</th>
<th>2nd Simulator Mean</th>
<th>P value 1st Simulator versus 2nd Simulator</th>
<th>1st Control Mean</th>
<th>2nd Control Mean</th>
<th>P value 2nd Simulator versus 2nd Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to complete intubation (min)</td>
<td>5.15 ± 1.63</td>
<td>0.88 ± 0.23</td>
<td>&lt; 0.001</td>
<td>6.74 ± 2.56</td>
<td>5.95 ± 1.92</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>No. of times hits mucosa</td>
<td>21.42 ± 10.10</td>
<td>3.00 ± 1.95</td>
<td>&lt; 0.001</td>
<td>25.8 ± 10.35</td>
<td>21.38 ± 11.62</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Time viewing mucosa (min)</td>
<td>2.24 ± 1.32</td>
<td>0.19 ± 0.18</td>
<td>&lt; 0.001</td>
<td>2.89 ± 1.76</td>
<td>2.73 ± 2.27</td>
<td>0.001</td>
</tr>
<tr>
<td>Percent of time viewing airway</td>
<td>58.54 ± 17.89</td>
<td>80.36 ± 14.67</td>
<td>0.004</td>
<td>58.8 ± 17.66</td>
<td>60.52 ± 23.36</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Values are mean ± s.d.

#### Quality assessments

- Question arises whether the study have enough participants to minimise the play of chance (control group only 8 participants)
- Randomisation claimed and described
- We are not very clear whether adequate adjustments been made for residual confounding. There may be other differences between the groups at entry to the trial. The researchers not produce a table listing key demographic information and other level of training (e.g. some experience participants may have tested this VR module before) separately for the groups.
- Outcome measurements were clearly described and determine in the same way between interventions
- Blinding of outcome measurements has not indicated
- Results answered the research question
- P value is reported where as confident intervals are unavailable

#### Overall conclusion

- Results of this research have shown that residents can be successfully taught the psychomotor skills of fiberoptic intubation in pediatric patients by using the Simulator.

**Purpose of the study**
- To investigate the operation performance after a VR training course

**Participants**
- Participants (n=12) were surgeons-in-training and novices in laparoscopic cholecystectomy.

**Comparison groups**
- Experimental group (n=12) consist of surgeons participating in VR training course and 12 other surgeons-in-training constituted the control group.

**Intervention**
- Open Xitact LS500 laparoscopy simulator platform (Xitact SA, Morges, Switzerland)

Structured questionnaires including multiple observation scales were used to assess performance.
**Outcome**
- Performance outcome parameters: “fluency,” carefulness, and “judgment” on part of the “clip-and-cut” procedure of the laparoscopic cholecystectomy

**Results**
- Observers judged the overall performance in the experimental group to be clearly superior.

**Quality assessments**
- The participants were not randomly assigned to either group.
- We are not very clear whether adequate adjustments been made for residual confounding. The groups were demographically comparable except for the number of laparoscopic cholecystectomies performed. Thus, it cannot be firmly stated that the participants were equally skilled at inclusion e.g. control subjects were in fact more experienced beforehand
- Effort was made to achieve blinding to minimise the observer bias.
- P value is reported where as confident intervals are unavailable

**Overall conclusion**
- This study showed that OR performance was significantly better in the experimental group
A number of studies have been undertaken to demonstrate that skills acquired during VR training transfer to the operating room. The background, results, and significance of these studies are reviewed in this article.

Authors pointed out that most of the cited VR-to-OR studies show skills transfer effects, it is necessary to look beyond these positive data, and to formulate practical recommendations pertaining to formative training of operative skills. As capabilities of VR simulation becomes feasible in future, ethical questions become more relevant and a new phase of study and validation has to be envisioned that defines additional methods to examine outcomes of training. This would require development and routine use of assessment systems for clinical performance.

Thus, the optimal use of new VR training platforms requires that the best possible assessment methods for the clinical OR be devised and validated. More intensive evaluation of this type could be used to guide implementation of innovative training methods such as VR, based on a dynamic process of continuous examination of performance, identification of performance outliers, and modeling of training activities to achieve carefully selected training goals based on expert performance behaviours.
Purpose of the study
- To demonstrate that virtual reality (VR) training transfers technical skills to the operating room (OR) environment.

Participants
- Sixteen surgical residents (11 male, 5 female) in postgraduate year (PGY) 1 to 4

Comparison groups
- Study group that would receive VR training in addition to the standard programmatic training (ST) and a control group that would receive ST only

Intervention
- Minimally Invasive Surgical Trainer-Virtual Reality (MIST VR) system (Mentice AB, Gothenburg, Sweden)
- It was run on a desktop PC (400-MHz Pentium II, 64-Mb RAM) with tasks viewed on a 17-inch CRT monitor positioned at operator eye level.
Outcome
Laparoscopic cholecystectomy procedural errors chosen as the study measurements

Results
- No significant differences were noted in visuospatial, perceptual, or psychomotor abilities between subjects randomized to ST and VR groups when assessed before the training phase of the study.

Figure 4. Total error number for each error type. LOP, lack of progress; GBI, gallbladder injury; LI, liver injury; intraperitoneal, incorrect plane of dissection; BNT, burn nontarget tissue; TT, tearing tissue; IOV, instrument out of view; AT, attending takeover. In all error categories except LI and TT, a greater number of errors were observed in the ST group than in the VR group.
Quality assessments

- Prospective, randomized, blinded study
- Blinding and randomisation stated but method not described
- Some doubts arise there were enough people in the study. Participants in each study arm not mentioned.
- The differences between the groups at baseline were not mentioned. This might, in part, account for any differences in outcome. Inclusion of confounding variables would have provided interesting comparison.
- Participants in each group followed up and data collected in the same way, reviewed at the same time intervals, in the same way. However, we are not sure whether there was loss-to-follow-up.
- Results answered the research question. P value is reported where as confident intervals are unavailable.

Overall conclusion

- The use of VR surgical simulation to reach specific target criteria significantly improved the OR performance of residents during laparoscopic cholecystectomy.
Purpose of the study
- To evaluate the difference between simulator and bedside training at an early stage of endoscopic training.

Participants
- 20 residents with no prior experience of performing endoscopy

Comparison groups
- Simulator and non-simulator groups

Intervention
- GI-Mentor II simulator (Simbionix USA Corp., Cleveland, OH, USA)

- Each subject performed endoscopy twice for assessment and performance was evaluated

Outcome
- Assessment of endoscopic skills. Eleven items were used in our criteria to evaluate the manipulation of the endoscope as well as accurate observation.

Results
- No significant difference in the total procedure time between the two groups.
- The scores of the items related to manipulation were significantly higher in the simulator group than in the non-simulator group.
- No significant differences between the two groups with regard to the items evaluating observational skills
Quality assessments
- Prospective randomized blinded study
- Blinding and randomisation stated and method described
- Other key differences between the groups at baseline were not mentioned. This might, in part, account for any differences in outcome
- Participants in each group followed up and data collected in the same way, reviewed at the same time intervals, in the same way. However, we are not sure whether there was loss-to-follow-up.
- Results answered the research question. P value is reported whereas confident intervals are unavailable
- Conclusion drawn supported by the study results

Overall conclusion
- The performance of endoscopy was improved by 5 h of simulator training.
- Authors concluded that the GI-Mentor II was useful for improving manipulative skills.

**Purpose of the study**
- To investigate the use of virtual reality simulator MIST-VR as an assessment tool in basic surgical skill course

**Participants**
- 13 Basic Surgical Trainees (10 men and three women) attending a Basic Surgical Skills course

**Control groups**
- Senior medical students (n = 13) with no laparoscopic experience and no formal training

**Intervention**
- PC-based MIST-VR (Minimally Invasive Surgical Trainer; Virtual Presence, London, SE1 2NL, UK)

```
Group A: Basic Surgical Trainees (n=13)  
Assessment with MIST-VR  
Basic Surgical Skills course  
Assessment with MIST-VR  
3 weeks and 3 months later  
Assessment with MIST-VR
```

```
Group B: Senior medical students (n=13)  
Four repetitions of the MIST-VR task
```

**Outcome**
- The time taken, (b) the number of movements made, and (c) the distance traveled by the instruments, (d) the economy of movement (i.e., the distance traveled by the instruments relative to the minimal distance possible), and (e) the number of errors made in completing the MIST-VR laparoscopic tasks
Results

### Table 1. Performance of trainees attending Basic Surgical Skills course (n = 13) (mean ± 95% CI)

<table>
<thead>
<tr>
<th>Repetition</th>
<th>Distance traveled (cm)</th>
<th>Distance ratio</th>
<th>Errors made</th>
<th>Movements made</th>
<th>Time taken (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1</td>
<td>35.6 (27.6–45.8)</td>
<td>4.0</td>
<td>9.4</td>
<td>27.5</td>
<td>30.3</td>
</tr>
<tr>
<td></td>
<td>(3.1–5.1)</td>
<td>(6.9–12.7)</td>
<td>(20.4–37.1)</td>
<td>(23.0–40.1)</td>
<td></td>
</tr>
<tr>
<td>Day 3</td>
<td>27.2</td>
<td>3.0*</td>
<td>3.8*</td>
<td>17.1b</td>
<td>19.0b</td>
</tr>
<tr>
<td>(n = 12)</td>
<td>(23.8–31.2)</td>
<td>(4.8–7.2)</td>
<td>(14.0–20.8)</td>
<td>(13.8–22.9)</td>
<td></td>
</tr>
<tr>
<td>3 wk</td>
<td>25.1a</td>
<td>2.8*</td>
<td>4.9c</td>
<td>12.9c</td>
<td>15.6c</td>
</tr>
<tr>
<td>(n = 12)</td>
<td>(21.8–28.9)</td>
<td>(4.2–5.7)</td>
<td>(10.3–16.0)</td>
<td>(13.2–18.5)</td>
<td></td>
</tr>
<tr>
<td>3 mo</td>
<td>29.3</td>
<td>3.2</td>
<td>6.7</td>
<td>16.0a</td>
<td>19.6d</td>
</tr>
<tr>
<td>(n = 11)</td>
<td>(24.6–34.9)</td>
<td>(5.3–8.4)</td>
<td>(11.8–21.8)</td>
<td>(15.8–24.3)</td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05
* p = 0.01
* p < 0.001

### Table 2. Performance of control group (n = 13) (mean ± 95% CI)

<table>
<thead>
<tr>
<th>Repetition</th>
<th>Distance traveled (cm)</th>
<th>Distance ratio</th>
<th>Errors made</th>
<th>Movements made</th>
<th>Time taken (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1</td>
<td>39.3 (28.2–54.9)</td>
<td>4.5</td>
<td>11.1</td>
<td>28.4</td>
<td>30.2</td>
</tr>
<tr>
<td></td>
<td>(3.2–6.4)</td>
<td>(8.1–15.2)</td>
<td>(20.9–38.5)</td>
<td>(22.8–39.8)</td>
<td></td>
</tr>
<tr>
<td>Day 3</td>
<td>35.5</td>
<td>4.0</td>
<td>8.3</td>
<td>20.2</td>
<td>22.8</td>
</tr>
<tr>
<td>(25.8–49.0)</td>
<td>(3.6–5.4)</td>
<td>(6.2–11.2)</td>
<td>(15.6–26.2)</td>
<td>(17.8–29.2)</td>
<td></td>
</tr>
<tr>
<td>3 wk</td>
<td>36.9</td>
<td>4.1</td>
<td>8.8</td>
<td>21.1</td>
<td>24.2</td>
</tr>
<tr>
<td>(26.2–51.9)</td>
<td>(3.6–5.7)</td>
<td>(6.2–12.4)</td>
<td>(15.4–28.9)</td>
<td>(18.1–32.2)</td>
<td></td>
</tr>
<tr>
<td>3 mo</td>
<td>31.9</td>
<td>3.5</td>
<td>7.8</td>
<td>17.6</td>
<td>20.6</td>
</tr>
<tr>
<td>(22.2–45.7)</td>
<td>(2.5–4.9)</td>
<td>(5.2–11.7)</td>
<td>(11.3–27.6)</td>
<td>(14.2–29.9)</td>
<td></td>
</tr>
</tbody>
</table>

Quality assessments

- Prospective case-control study
- Blinding and randomisation not stated.
- Differences between the groups at baseline were mentioned.
- Participants in each group followed up and data collected in the same way, reviewed at the same time intervals, in the same way. However, there was loss-to-follow-up.
- Results answered the research question. P value and confidence intervals are available.
- Conclusion drawn supported by the study results

Overall conclusion

- Basic Surgical Skills course produces an improvement basic laparoscopic skill which is quantifiable using the virtual reality simulator MIST-VR.

**Purpose of the study**
- To determine whether training on a VR simulator led to the transfer of skills to operation room

**Participants**
- First- and second-year surgical trainees (n=20)

**Intervention**
- In this study the SIMENDO VR simulator (DeltaTech, Delft, the Netherlands) was used.

**Comparison groups**
- Control group did not receive any further manual VR training after the basic skill training on VR simulator. Instead they viewed three consecutive video demonstrations
### Outcome
- Objective analysis parameters were: time taken to tie the knot and number of predefined errors made.
- Subjective assessments were also made by two laparoscopic surgeons using a global rating list with a five-point Likert scale.

### Results

<table>
<thead>
<tr>
<th></th>
<th>Group A (experimental), n = 10</th>
<th>Group B (control), n = 10</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total error score</strong></td>
<td>24 (10–40)</td>
<td>36 (17–54)</td>
<td>0.030</td>
</tr>
<tr>
<td><strong>Total time taken to drive the needle through the tissue (seconds)</strong></td>
<td>118 (60–510)</td>
<td>203 (70–647)</td>
<td>0.253</td>
</tr>
<tr>
<td><strong>Total time taken to tie knot (seconds)</strong></td>
<td>262 (69–406)</td>
<td>374 (169–600)</td>
<td>0.034</td>
</tr>
</tbody>
</table>

### Quality assessments
- Prospective case-control study
- Blinding and randomisation stated and method described.
- Differences between the groups at baseline were not mentioned. This might, in part, account for any differences in outcome.
- Participants in each group followed up and data collected in the same way, reviewed at the same time intervals, in the same way. However, there was loss-to-follow-up.
- Results answered the research question. P value is reported where as confident intervals are unavailable.
- Conclusion drawn supported by the study results

### Overall conclusion
- Results showed surgical trainees who received knot-tying training on the VR simulator were faster and made fewer errors than the controls.
- Authors concluded that the VR module is a useful tool to train laparoscopic knot-tying.