Training and Assessment of Basic Laparoscopic Skills
- Development of an Evidence-Based Simulation
Curriculum

Cand.med.
Anders Johan Debes

Ostfold Hospital Trust,
Department of Surgery
Fredrikstad, Norway

and

University of Oslo, Norway
2013

ØSTFOLD
HOSPITAL TRUST

**UNIVERSITY
OF OSLO**

© Anders Johan Debes, 2014

*Series of dissertations submitted to the
Faculty of Medicine, University of Oslo
No. 1930*

ISBN 978-82-8264-841-7

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinssen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika Publishing.
The thesis is produced by Akademika Publishing merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Table of contents

LIST OF ORIGINAL PAPERS	5
ACKNOWLEDGEMENTS	7
1.0 INTRODUCTION.....	11
1.1 «SEE ONE, DO ONE, TEACH ONE»	13
1.2 SHORT HISTORY OF LAPAROSCOPIC SURGERY.....	16
1.3 SURGICAL SIMULATION – A LESSON FROM AVIATION	18
1.4 LAPAROSCOPIC SIMULATORS	20
1.5 SURGICAL COMPETENCE AND CORE LAPAROSCOPIC COMPETENCE.....	26
1.6 EDUCATIONAL THEORY – A SHORT INTRODUCTION	29
1.7 LEARNING CURVES	35
1.8 OBJECTIVE ASSESSMENT OF SURGICAL COMPETENCE.....	37
1.9 VALIDATION OF SURGICAL SIMULATORS	43
1.10 RELIABILITY	46
1.11 LAPAROSCOPIC BASIC SKILLS CURRICULA.....	47
2.0 GENERAL AIM OF THE STUDY	51
2.1 SPECIFIC AIMS	51
3.0 MATERIALS AND METHODS.....	53
3.1 SIMULATORS USED IN THE STUDIES	53
3.2 ICSAD/ROVIMAS – ELECTROMAGNETIC TRACKING OF HAND MOTION.....	58
3.3 DATA ACCESS AND ETHICAL CONSIDERATIONS	60
3.4 STATISTICS	60
4.0 SYNOPSIS OF PAPERS AND INDIVIDUAL MAIN CONCLUSIONS.....	63
4.1 PAPER I.....	63
4.2 PAPER II	65
4.3 PAPER III.....	67
4.4 PAPER IV.....	69
5.0 DISCUSSION.....	71
5.1 METHODOLOGICAL CONSIDERATIONS	71
5.2 FINDINGS AND IMPLICATIONS.....	77
6.0 FUTURE PERSPECTIVES.....	91
REFERENCES	94

List of original papers

- I. Debes AJ, Størkson RH, Jacobsen MB. **Curative rectal cancer surgery in a low-volume hospital: A quality assessment.** *Eur J Surg Oncol.* 2008 Apr; 34(4): 382-389.
- II. Debes AJ, Aggarwal R, Balasundaram I, Jacobsen MB. **A tale of two trainers: virtual reality versus a video trainer for acquisition of basic laparoscopic skills.** *Am J Surg.* 2010 Jun; 199(6): 840-5.
- III. Debes AJ, Aggarwal R, Balasundaram I, Jacobsen MB. **Construct, content and face validity of D-Box; a web-cam based laparoscopic basic skills box-trainer.** *Submitted.*
- IV. Debes AJ, Aggarwal R, Balasundaram I, Jacobsen MB. **Construction of an evidence-based, graduated training curriculum for D-Box, a webcam-based laparoscopic basic skills trainer box.** *Am J Surg.* 2012 Jun; 203(6): 768-75.

Acknowledgements

This PhD thesis initially started as a quality improvement project within the Department of Surgery, Sykehuset Østfold. However, it quickly transformed into a PhD-project in laparoscopic training by using simulators, a subject neither my supervisor nor I had any previous experience with. I would therefore like to express my sincere and deepest gratitude towards all those who have, in one way or another, helped me initiate, perform and complete these studies:

- My principal supervisor **Morten Bj. Jacobsen** for recruiting me into the field of research and making it possible for me to start, and complete, this PhD. I would also like to thank him for always being available for discussions, for his wide span of scientific enthusiasm and continuous support.
- Colleague/friend/co-author **Ragnhild Størkson** who recruited me to participate in her planned study on laparoscopic colonic resections. It has been a great pleasure to work parallel to her, sharing the same joys and frustrations of doing a PhD, attending the same PhD-courses until we finally ended up submitting our theses on the same day.
- My secondary supervisor **Rajesh Aggarwal** and his former colleague at Imperial College, London, **Indran Balasundaram** for providing experienced knowledge, for cooperation in designing studies, technical support and for co-authoring on three of the manuscripts.
- **Teodor Grantcharov** for inspiration and expertise on the use of surgical simulators in the initial phase of these studies, and for introducing me to many influential people in this field.

- Several inspirational colleagues deserve a honourable mention; **Arne Rosseland, Ole Christian Olsen, Fredrik H. Halvorsen** and **Ronald Mårvik**. Thank you for your devotion and passion for training surgeons and making surgery safer for all.
- I am also grateful to **Magdalena Chmarra** and **Cecilie Våpenstad** for their detailed comments, suggestions and helpful points when writing this thesis.
- My two employers during this time period, Department of Surgery, **Sykehuset Østfold** and Department of Urology, **Akershus universitetssykehus**, for providing time and financial support, making it possible for me to finish my thesis.
- All participating surgeons, interns and medical students for contributing with their time, effort and enthusiasm.
- All my colleagues for covering the shifts and doing the extra work when I had time off for writing, for continuous support and for making my days better.
- My dear family, especially my mother **Mona**, for all the loving support and for bringing me up to believe that hard work, passion and determination is essential to reach your goals.
- My friends for not giving up on me, for not killing me totally with the question: *When will your thesis be finished?* and for calling or taking me out, making sure I had at least some social training during the most intense periods working with this thesis.
- Last, but above all, I want to thank my partner **Arne**, for his continued patience, inspiration and encouragement throughout this whole process. He has been there to witness my highs and lows; always supportive and constructive, sacrificing weekends and holidays so that I could become a sub-specialised surgeon and also complete this thesis at the same time. I will forever be grateful.

Abbreviations

AR	Augmented reality
CRM	Crew resource management
EOM	Economy of movement
EWTD	European Work Time Directive
GRS	Global ratings scale
KBB	Knowledge-based behaviour
MAP	Motion analysis parameter
MIS	Minimal invasive surgery
MIST-VR®	Minimal Invasive Surgical Trainer
OR	Operating room
OSATS	Objective structured assessment of technical skills
RBB	Rule-based behaviour
SBB	Skills-based behaviour
THM	Total number of hand movements
TPL	Total path length
VAK	Visual-auditory-kinaesthetic
VR	Virtual reality

1.0 Introduction

In the last few decades, surgery has changed profoundly by the introduction of minimally invasive surgery (MIS, e.g. laparoscopy) (1). Procedures, such as cholecystectomies, that previously involved long skin incisions and a hospital stay of 7-14 days, are now routinely performed as day-case procedures with only a few cm-sized skin incisions (2). Compared to open surgery, the patients undergoing a laparoscopic procedure experience less postoperative pain, fewer adhesions, more acceptable cosmetic result and a shorter time before they can resume previous daily activities (1,3,4). Laparoscopic surgery has also resulted in shorter period of hospital stays. As a consequence, the number of patients to be treated can be increased, hence reducing the cost per patient (4-6). Moreover, laparoscopy in children and elderly patients improves outcomes and results in fewer complications in comparison to open surgery (7,8).

Additional development of advanced surgical equipment like mechanical stapler devices, the use of laser therapies and electrosurgical/ultrasound based dissecting tools, in combination with high definition video equipment, has provided the surgeons of today with tools that offer a higher quality of operative care and level of outcome than before. The recent development and subsequent implementation of robotic surgery has further increased the possibilities for complex surgery. Logically, whilst surgery is getting increasingly more advanced, surgeons face greater challenges, and the risk of medical errors is higher than before.

Initially, introduction of laparoscopic surgery resulted in a higher number of complications, thus shedding a bad light on the technique (9). Research and empirical knowledge over time attributed this to lack of sufficient training (realisation of a longer learning curve for laparoscopic surgery compared to open surgery). This led to the acknowledgement of necessity for specific training for laparoscopic skills (10,11).

About one in ten patients experiences unintended injuries – adverse events –during his/her hospital stay. Adverse events lead to complications that can result in prolonged hospitalisation, disability or death caused by the actions of healthcare professionals rather than by effects of underlying medical condition of a patient (12). 50-70% of adverse events occur during surgical interventions (12). A recent national hospital safety campaign, initiated by the Norwegian health authorities, shows that 16% of hospitalised patients in Norway experience adverse events (13). Annually more patients are killed by hospital-related adverse events than breast cancer and AIDS combined (14,15). Systematic reviews estimate that 40-50% of such events are preventable (12,14). A number of individual or complex factors influence the outcome after surgery (16). However, over 40% of adverse events occur in the operating room (OR), indicating that surgical skills are one of the most important single factors (12).

Improving patient safety has received great interest from the surgical community as well as health inspectorates, the media, patients and insurance companies worldwide (17-19). Improving patient safety entails improving quality of care for surgical patients. Although there will never be a time when surgery is devoid of risk, the quest for minimizing the risk as much as possible must continue. Securing adequate surgical

education and training is crucial. Herein lies the challenge for surgical programme directors to design evidence-based, time-efficient, cost-efficient and feasible training programmes that ensure sufficient surgical competence for surgeons in training today and for the future.

By implementing simulators into the surgical curricula, program directors gain access to valuable tools for training and objectively assessing surgical skills beyond mere procedural volume and subjective evaluation by supervisors. The inspiration for commencing this thesis grew from my work on the internal review of outcome after curative rectal surgery, where I discovered the problems associated with defining what constitutes «a good surgeon» or «surgical proficiency».

1.1 «See one, do one, teach one»

A clear paradigm shift in surgical education was introduced in 1904 when William S. Halsted published his ground breaking paper «*The training of the surgeon*», describing his structured educational system (20). This is the first documented surgical training programme in modern medicine and was based on the slogan «*See one, do one, teach one*». This has since been the mantra of surgical education for many years. Until now, the OR has been the classroom, and the transfer of knowledge and skills in surgery has followed a master-apprenticeship model in a predominately hands-on manner.

The educational path for becoming a sub-specialised surgeon in Norway typically takes 8-9 years after internship (which lasts for 1,5 years after graduation). Specialisation is a combination of 5-6 years of general surgery followed by three years of sub-

specialisation (21,22). When a candidate applies for board certification for his/her speciality or sub-speciality in Norway, there are three requirements that must be fulfilled:

1. **Time:** The candidate must provide signed documentation of the duration and contents of each placement.
2. **Knowledge:** The candidate must provide copy of all course diplomas from the list of compulsory courses required.
3. **Logbook:** The candidate must document that he/she has performed (independently or as the main surgeon under supervision by a specialist) a minimal number of operations on a specified list of surgical procedures.

Today, none of the requirements involve any objective assessment of the candidates' skills during, or at completion of his/her specialisation in Norway. Only at the conclusion of a few of the compulsory courses, multiple-choice tests for knowledge are administered. There are no practical examinations administered to objectively assess the technical skills of the candidates during their specialisation. At present, no written or oral examinations of the candidates are administered at completion of the specialisation period before board certification is granted (23). The whole board certification process is thus based on an assumption that the necessary surgical training is acquired within the time requirements, and that the necessary theoretical knowledge is covered by the compulsory courses and by self-studies and organized tutoring within the department/hospital. This represents a time-based approach. A proficiency-based

approach for progression and certification, on the other hand, seems to be individualised, more efficient and thus a logical approach.

Procedural numbers are used in many settings to define thresholds for granting privileges or accreditation status. They are often used as a predictor for level of quality or outcome, however, for this they will always just be surrogate measures. The relationship between hospital (or surgeon) caseload and outcome is not straightforward. In the case of rectal cancer surgery, where several studies indicate that high volume predicts better outcome (less postoperative morbidity, mortality and local recurrences), there are still significant variations in outcome (24-26). There can be many factors contributing to this, but surgical skills definitely plays a part, when degree of specialisation of surgeons vs. outcome shows a stronger correlation than high-volume surgeons vs. outcome (24,27).

The time-based approach is a remainder from a time when surgeons worked more than 120 hours per week, when the sheer time spent at the hospital, on call and in the operating theatre, provided sufficient operative caseload to produce adequately trained surgeons. Introduction of the European Work Time Directive (EWTD) has effectively reduced the number of hours a surgeon in training is spending in the OR (28). A post-EWTD-study in Britain (29) demonstrated that the operative hours of a surgical trainee were cut in half over the last decade. A comparable trend can be observed in the USA (30-32), especially at the junior levels, but also for members of the surgical faculty leading to concerns with sufficient supervision (33). In Norway, this has had a lower impact, probably because of already existing strict work-time legislations. However, the

trend is evident in Norway as well; an average workweek for a general surgeon in 2000 was 44.9, and in 2010 was 42.6 hours¹.

In addition to reduced time in the OR, other factors are threatening the possibility of surgeons-in-training achieving the level of competency required, to provide safe and high-quality surgical treatment, both for today and in the future. In healthcare now, there is a greater demand for increased «production» and decreased costs. This results in a trend towards a shift in focal point from providing high quality healthcare and education of surgeons towards economic goals like budgets and cost-per-patient (34). There is also an increased demand for administrative reports and documentation for legal purposes, and a reduction in support staff leaves the surgeons spending more time handling miscellaneous paperwork. All together, these factors are further reducing the surgeons' hours in the OR and, thus, posing both an immediate and imminent threat to developing sufficient surgical competency.

1.2 Short history of laparoscopic surgery

Laparoscopy represents a relatively new approach in surgery. As many other medical discoveries, it is difficult to credit one individual for pioneering laparoscopy. However, in 1901 Georg Kelling (1866-1945) performed the first laparoscopic procedure using a cystoscope in a dog's abdomen and in 1910 the swede Hans Christian Jacobaeus (1879-1937) reported the first laparoscopic operation in humans (35,36).

¹ Personal communication: Anders Taraldset, The Norwegian Medical Association

The method gained very little enthusiasm given its supposed limited application and was for many decades used mainly for purposes of diagnosis and performance of simple procedures in gynaecology. Gynaecologists Hans-Joachim Lindemann (1920-2012) and Kurt Semm (1927-2003) from the Universitäts Frauenklinik in Kiel, Germany, performed CO² hysteroscopy during the mid-seventies. After the development of an automatic CO²-insufflator (1963), the thermo coagulation device (1973) and, finally, an electronic insufflator, the advent of the laparoscopic approach was firmly established. In 1980, Dr. Semm performed the first laparoscopic appendectomy (37). Later on, he continued innovative research, developing instruments and extending his spectrum of laparoscopic procedures. He also showed a great interest in teaching his methods; in 1985, he constructed the Pelvi-trainer (= laparo-trainer) – an inanimate surgical simulator whereby colleagues could practice laparoscopic techniques.

The introduction of computer chip video camera in the late 1980s was a pivotal event in the field of laparoscopy. This innovation provided the means to project a magnified view of the operative field onto a monitor freeing both of the surgeon's hands and facilitating performance of complex laparoscopic procedures. In 1987, in Lyon, France, the first laparoscopic cholecystectomy was credited to dr. Philippe Mouret (1937-2008) (38). Prior to this, the surgical community had shown a great scepticism for the laparoscopic approach, but the method gained increasing acceptance and, within 5 years, the laparoscopic cholecystectomy was considered the «best approach». The first World Congress on Surgical Endoscopy was held in Berlin in June 1988, assembling a group of approximately 500 international experts in surgical endoscopy.

Since then, developments in technology have brought new, previously inconceivable methods, tools and possibilities for the surgeons of today. The rate of these significant developments and the on-going technological innovations also generate a perpetual need for improving present and learning new skills. This paradigm-shift in disruptive technology needs to be followed by a consequent paradigm-shift in surgical education.

1.3 Surgical simulation – a lesson from aviation

Simulation technology in surgery was initially adopted from the aviation community where it has been utilized both to train individuals, like pilots, navigators and flight attendants, as well as full teams in «crew management training» (39). A number of serious airplane accidents in the seventies, and especially the plane crash on the island of Tenerife in 1977, killing 583 passengers, raised concerns about human performance and errors in advanced technologic environments (40). The specific accident involved two Boeing 747 aircrafts crashing whilst taxiing and taking-off on a foggy runway. The following investigation concluded that the cause was multi-factorial, but mainly caused by non-standard communication and misunderstandings between the air traffic controllers and pilots of the two airplanes – so-called «human errors».

Aviation authorities worldwide made substantial changes to airline and aircraft regulations as a direct consequence of the Tenerife-accident (40). Cockpit procedures were consequently changed, safety and procedural checklists were introduced, hierarchical relations among the crewmembers were played down, and more emphasis was placed on team decision-making by mutual agreement. Crew Resource Management (CRM) became the name of the systematic approach that was initiated

(39). Considerable interest in this area resulted in a number of research projects, investment of both economic and scientific resources leading to the development of advanced simulators for use in the aviation industry. Today, a pilot can learn, practice and become skilled at all aspects of a flight, including how to deal with unexpected events, without even leaving the ground. Following the example from the aviation industry, other fields of expertise, where advanced technology, human interface and a non-tolerance for errors like nuclear plants and the Military, have embraced simulation technology and incorporated it into their training and certification procedures (39,40).

Recently, the surgical community also has adopted another important safety effort from the aviation industry: checklists. The implementation of the 19-item surgical checklists, developed after the «Safe Surgery Save Lives» campaign initiated by the World Health Organisation (WHO) in 2008, demonstrated significant reduction in death rates and complications after non-cardiac surgery in adults (41). These checklists are now being used prior to all surgical procedures in an increasing number of Norwegian hospitals.

Standard training and assessment requirements have been utilised for testing pilot aptitude within the aviation industry for many years. These requirements also institute the threshold for pilot progression from one grade to another, transfer from one aircraft type to the other, and continuous competence assurance (39). Selection of future pilots is based on a series of tests, designed to evaluate all of the potential pilot's abilities assumed to be relevant for becoming a proficient and safe pilot. These tests involve assessment of intelligence, tests for spatial orientation, mental capacity (for processing and acting on multiple inputs) and mental agility (42,43). Combining the results of these

tests with a weighted appraisal of the applicant's age, previous flying experience and internal motivation, a probability score is estimated and aids in prediction of success or failure in pilot training.

In the late seventies and early eighties, reliable methods to assess and predict success/failure of candidates for surgical training were lacking. Developing assessment tools to aid in the selection of the «most promising», and thus time and cost-effective, surgical trainees and the pursuit for error reduction (44), became the major driving forces for the initial development of surgical simulators (45).

1.4 Laparoscopic simulators

The hunt for educational tools for training laparoscopic skills in a safe, reproducible and standardised way, led to the development of surgical simulators. The image-guided nature of laparoscopic surgery makes it suitable for simulation training. Today, after three decades of on-going development, numerous laparoscopic simulators are commercially available (46,47) and extensive work has been done to validate them in order to implement them into surgical training curricula (46,48-51) (validation of simulators is further described in section *1.9 Validation of surgical simulators*).

The concept of «surgical simulation» spans wide and includes everything from practicing sutures on small plastic patches to removing moles all the way to advanced, computerised laparoscopic or endovascular simulators that allow full procedural. A laparoscopic simulator may vary from a simple homemade setup consisting of a shoebox with an old, discarded laparoscope – to highly advanced, virtual reality (VR)

simulators with realistic haptic feedback, objective assessment and multimedia-enhanced training modules (47). Currently, there are two categories of surgical laparoscopic simulators; box-trainers and computer based, VR-simulators. There are also emerging simulators combining the advantages of the VR-simulators and box-trainers – so-called hybrid simulators, or augmented reality (AR) simulators (52,53). Each simulator has specific inherent properties, which provides the user with a different set of benefits and disadvantages. Laparoscopic simulators are used for training and assessment of both basic skills, advanced skills and procedural training. Advanced skills and procedural training is beyond the scope of this thesis, therefore, only simulators for training basic (psychomotor) laparoscopic skills are further described and discussed.

Box-trainers, also called video-trainers, mechanical simulators, trainer boxes or inanimate trainers, are available in many different designs and provide a wide variation of tasks. Usually, they involve either the use of animal organs or inanimate objects being manipulated by real, commercially available laparoscopic instruments. By including real laparoscopic instruments and physical objects or animal tissue (organs) for the tasks, tactility is represented at the same level as in the OR. For imaging, these simulators typically use a spare laparoscopic camera, a web-cam or small video camera, producing an image on a monitor. Box-trainers are used for training basic laparoscopic skills, suturing skills and to some extent also procedural training (for example laparoscopic cholecystectomy on cadaver pig organs).

Figure 1 - FLS (Fundamentals of Laparoscopic Surgery)

VR-based simulators usually are used to train basic laparoscopic skills (similar to the box-trainers), but also suturing skills, complex tasks and procedural tasks (46).

Regarding the instruments, the handles are often realistically designed, but the distal («intra-abdominal», functional) part of the instrument is only digitally presented on the screen. The simulator software usually contains realistic backdrop images and can render an array of instruments allowing instrument-changes during training. Dependent on its nature, the tasks can be presented in front of an abstract or realistic scene/backdrop. Given the digital representation only of the instruments' tips, VR-simulators lack the natural haptic feedback and, therefore, some of them try to recreate it with additional hardware. To provide realistic haptic sensation is, however, technically challenging and the available VR-simulators today are still deficient in emulating the natural feel of haptic feedback.

Figure 2 - MIST-VR is an example of a VR-simulator

Assessment and feedback on performance are essential parts of a learning process.

Where box-trainers usually offer no objective assessment other than time-to-completion, VR-simulators typically offer objective assessment of multiple variables along with teaching and guiding resources (46).

Hybrid simulators combine the physical aspects of a box-trainer with the ability to track instruments for assessment purposes (54). The result is a simulator that provides natural haptic feedback coupled with objective assessment tools. Hybrid simulators usually contain a multimedia learning module and real-time guidance, including also objective feedback to some degree.

Figure 3 – ProMIS® is a hybrid laparoscopic simulator

Both VR-based simulators and box-trainers have shown to provide valuable training for basic laparoscopic skills (55-60). There is also growing evidence that training with laparoscopic simulators is efficient for developing skills that lead to improved performance in real operations (57,61,62). Bridges and Diamond published in 1999 a register-based study on the costs of training surgical residents in the OR (63). They included almost 15.000 procedures from 62 different procedure categories (open and laparoscopic approaches) performed by over 1000 general surgical residents. Although methodological issues can be raised, they recommend increased use of digital modalities and simulation technology for basic skills training outside the OR because of increased OR-time (and thus increased costs) when teaching/supervising during procedures (63). These calculations are optimistic, they were based exclusively on the extra time spent in the OR training junior surgeons, and did not even include the costs of possible complications and errors during training.

Training on simulators, outside the OR, can provide several advantages. The most important ones are:

- **Patient safety** – basic surgical skills are acquired without any risk to patients.
- **Ethical considerations** – possibility to minimize training on animals and patients.
- **Repeatability** – practice of specific parts of an operation can be done any time, without the need to wait for a patient that requires that special procedure. Thus, it is possible to generate more effective volume training in a shorter amount of time.
- **Assessment** – possibility to objectively assess skills of the candidates to assure pre-trained, qualified surgeons before entering the OR. The assessment will also act as an documentation of skills.
- **Self-training/availability** – training can be done without formal supervision, and when the candidates have spare time during clinical hours.

Currently, the ideal simulator or curriculum for training novice laparoscopic surgeons has not yet been established, and studies comparing box-trainers and VR-simulators show contradictory results (48,60,64-66). Hence, further research is needed before we know what the optimal application of simulators is and before «perfect» simulators are available on the market. The difference in purchase costs between a box-trainer (\$5.000) and a VR-simulator (starting at approximately \$50.000 without haptic feedback) is in the tenfold-ratio and also raises important issues about cost-efficiency.

Several studies have demonstrated that both types of simulators provide acquisition and retention of skills when training for MIS, but no consensus is established on what specific training program is necessary for having a significant positive impact on surgical performance in the OR (67).

1.5 Surgical competence and core laparoscopic competence

All patients and next of kin want a competent surgeon when a procedure is necessary. However, measuring surgical competence is a difficult mission. Any discussion on competence must therefore start with a definition. This is however also an elusive task. New Webster's Dictionary of the English Language (1981) reads: *«competent adj. (Fr. competent, competer, to be sufficient; L. compete, to be suitable). Answering all requirements; suitable; fit; adequate; having legal capacity or power; rightfully or lawfully belonging. Competence; state of being competent; adequacy; sufficiency;... »* (68). This definition takes competence beyond just abilities and skills, but includes also the legal aspects, as in «legally qualified to perform an act». Surgical competence is generally not defined as broad as this. The American Board of Medical Specialities (ABMS) and The Accreditation Council for Graduate Medical Education (ACGME) have defined a set of criteria that define competence in medicine, which includes six components (69):

- *Knowledge*
- *Patient care*
- *Interpersonal and communication skills*
- *Professionalism*
- *Practice-based learning*
- *Improvement and system-based practice*

Competence is multifactorial in nature, with knowledge, judgement, behaviour and technical abilities each playing a key role (70,71). However, what differentiate a competent surgeon from other competent non-surgeon specialists are the specific manual skills necessary to perform surgical procedures (72).

A surgeon never works alone in the OR, and he/she cannot face per-operative complications or crisis-situations without a team of other healthcare providers. In surgery, as well as in aviation, competency exists simultaneously within the team and on an individual level. Both levels are equally important and contribute synergetic toward achieving the best possible «product» or outcome. The team competence is represented by the total cognitive knowledge, non-technical skills (including leadership, communication skills and decision-making) and psychomotor skills of all the team-members combined. Individual competence is the knowledge, non-technical skills and psychometric skills represented by each team-member. Simulation technology is suitable to practice both levels of competence. However, training in teams requires a different setup and array of assessment methods and is therefore beyond the limitations of this thesis.

Although surgical competence is multimodal in nature, proficiency in the specific laparoscopic dexterity skills is crucial to obtain a successful result. The laparoscopic technique challenges the surgeon with a set of very specific psychometric tasks, combined with an altered visual input and tactile sensory feedback (73-75). The three-dimensional operative field, known from open surgery, is displayed as a two-dimensional image (often with image quality degradation) on a monitor and the lack of a binocular image reduce the depth-of-field information, creating an eye-hand coordination challenge. The tactile feedback is also altered, mainly by the fulcrum effect of the instruments passing through the abdominal wall, the elongated instruments, friction between the instruments and trocars and the reduced degrees-of-freedom (DOFs) from six to four (55,56,76,77). Laparoscopic surgery moreover favours surgeons with developed ambidextrous skills and a greater sense of spatial relationships (78). All of these hindrances are laparoscopy-specific and represent an increased technical challenge. The specific skills required also separate the training for laparoscopic operating technique from the way open surgery is taught. Laparoscopy is not as intuitive as open surgery and, therefore cannot, and should not be taught in the same way as techniques in open surgery. The image-guided nature of laparoscopic surgery leaves the trainee to observe all movements of the distal part of the instruments on a monitor, without simultaneously observe how the surgeon manipulates the instruments from the outside. This is different compared to how one can observe and learn an open procedure. Acquiring the specific abilities to overcome the psychometric and sensory challenges in laparoscopy is essential and is the main focus of laparoscopic basic skills training (74,79,80). Only when the basic skills are practiced to the extent

that the surgeon does not have to actively think and plan before every move, when the manipulations of the instruments are smooth, performed sub-conscious and reflex-like, can he/she focus on learning the individual steps of the procedure itself.

1.6 Educational theory – a short introduction

As mentioned, performing surgery requires a broad spectre of skills involving not just the technical skills, but also surgical knowledge, abilities to lead, abilities to work in a team and good decision-making skills (70). Learning surgical procedures involves teaching adult professionals new skills. The gold standard of surgical training has traditionally been sheer volume-based, supervised training, following the Halsteadian master-apprentice model. Given the lack of training opportunities, mainly because of reduced supervised time in the operating theatre, there has been an increased interest for simulation training using different modalities. This new pathway of surgical education is based on established educational theories for acquiring psychomotor skills to expertise.

1.6.1 Learning styles

People are cognitively different and consequently have different preferred learning styles. For simplification one can divide learning styles into three categories called the VAK-styles (visual-auditory-kinaesthetic), which is supported by most researchers in this field (81). It provides a systematic view of people's preferred, or dominant, learning style, which draws on the individual's strength and personality.

- **Visual learners** experience the most effective learning by absorbing written information, pictures, diagrams, films etc. Visual learners make up about 65% of the population (71).
- **Auditory learners** learn most effectively by listening to spoken words, sounds, noises etc. Auditory learners make up about 30% of the population (71).
- **Kinaesthetic learners** require physical experience like touching, feeling, holding, moving etc. for the most effective learning. They are uncommon and make up about 5% of the population (71).

1.6.2 Learning cycles

Several models/theories of psychomotor skills learning exists in the literature of other disciplines (71). Some are highly applicable to learning practical skills by simulation, and these can successively be extrapolated into the world of surgery. One model that incorporates a further understanding of the individual learning styles and also describes the whole learning process elegantly is David Kolb's Experiential Learning Theory (ELT), first published in 1984 (82). The original theory describes four separate learning styles that relate to a four-stage learning cycle (Figure 4). The cycle is based on the assumption that people learn in a continued (cyclic) manner, through four connected stages. When the cycle is closed, all four stages have been addressed, and an effective learning process has taken place.

Figure 4 - Kolb's learning cycle

Immediate and concrete experiences lead to observations and reflection. These reflections are then absorbed, translated and lead to experimental actions, which again lead to new concrete experiences. Whatever influences a person's preferred mode of learning, the learning style itself is a product of two pairs of variables according to Kolb (82). Kolb arranges them as two crossing continuums, like axes in a diagram (Active Experimentation vs. Reflective Observation and Concrete Experience vs. Abstract Conceptualisation). One axis is the «thinking-axis» and the other the «processing-axis». Kolb then postulates that the endpoints of the axis represent conflicting learning styles and that the learner actively chooses how to react when faced with a new learning experience. The learner internally decides whether he/she wants to «do» or to «watch»

and at the same time «think» or «feel». The resulting four products represent the four preferred learning styles:

- **diverging** (feeling and watching)
- **assimilating** (thinking and watching)
- **converging** (thinking and doing)
- **accommodating** (feeling and doing)

Even though Kolb thought of these stages as a continuum that one follows, usually people have a natural, subconsciously preferred learning style. It is therefore necessary that programme directors keep this in mind when designing training programmes. The offered teaching style should ideally match the preferred learning style of the individual trainee(s).

1.6.3 Learning at different levels of behaviour

Translational research on human behaviour from other disciplines suggests that effective learning in laparoscopy should be adapted to the level of human behaviour (71,83). Three distinct levels of human behaviour have been defined:

- **Skill-based behaviour** (SBB) represents the skills performed without conscious control, in a highly automated fashion, using fast motor programmes for selection of appropriate muscles to control. These motor programmes are based on an accurate internal representation of the task, the system dynamics, and the environment at hand. An example of an everyday activity that requires skill-

based behaviour is walking, or brushing your teeth. Many tasks in laparoscopy can be considered a sequence of skilled acts like grasping, pulling, cutting or dissecting. An experienced surgeon performing a suturing task will perform this smoothly, without conscious thoughts behind every individual movement. At this stage the mental workload of the trainee is reduced and enables him/her to focus on the decision-making, rather than the execution of every movement.

- **Rule-based behaviour (RBB)** is the next level of human behaviour. At this level, skill performance is based on stored rules and/or procedures. These rules may be derived from the performer's own empirical experiences, from another person's expertise, or from books. The performer's previous success and experience leads to selection of the appropriate rule or procedure. An example of RBB is when performing a laparoscopic cholecystectomy; a complete dissection of Calot's triangle triggers the rule that the cystic artery and duct may be clipped next.
- **Knowledge-based behaviour (KBB)** describes the necessary behaviour where unfamiliar situations occur and no rules are available, like when encountering perioperative complications. At this level, information is perceived as symbols, i.e. collections of conceptual information. When experiencing a perioperative complication, the surgeon already has a goal or plan prior to the operation, but something happens that requires a different approach. The surgeon must then make an overall analysis of the situation and mentally develop multiple alternative plans, which then can be tested against the previously set goal. Now, the surgeon can choose the best strategy for counteracting the complication and

accordingly execute the plan. Serious complications that occur during MIS require a great deal of knowledge-based behaviour of the surgeon.

This model indicates further that training must also take place at different levels, and that different training modalities are required for each level. For the low-level (SBB), training simulators aid in learning basic skills, such as manipulating the instruments and camera. For higher levels (RBB and KBB), more sophisticated methods must be applied. When training for «how to deal with unexpected events», such as bleeding during operations, or even instrument breakdowns or power failure, highly advanced, VR-based interactive simulators could be developed. This would enable the trainees to sharpen their responses and decision-making skills in a safe environment.

1.6.4 Learning stages

Fitts and Posner's three-stage theory of motor skills acquisition was published in 1967 and has since gained wide acceptance in the literature of motor skills and surgery (84). In short, their theory comprise of the cognitive, integrative and autonomous stages of learning a new task/skill. The skills are acquired in a successive manner:

1. **Cognitive stage** – (also called the observation and imitation stage) during which the trainee intellectualises the task, and identifies the component parts of the skill. Each task is performed in an erratic way, and in obvious distinctive steps, without flow of movement.
2. **Integrative stage** – where the different component parts of the skill/task is (by repetitive training) becoming smoother and translated into one

motor skill. Performance is more fluent, but the trainee must still concentrate and think on how to execute the task.

3. **Autonomous stage** – where the task is performed more or less automatically without the trainee's conscious thought or attention to execute the task. Movement is smooth and the trainee can move on and focus on other parts of the procedure. Not all trainees reach this stage.

The learning stages can be illustrated by looking at a trainee who learns how to suture. First he/she must learn how to hold the instruments, secure the needle, place the stitch, do the throws and tighten the knot, before finally cutting the suture. In the beginning this is a deliberate, conscious process during which the trainee thinks ahead and plans every move. Next, repetition and objective feedback enables the trainee to synthesize the different components of the skill into a fluid, automated suturing movement. In essence, Fitts & Posner's principles propose a sequential learning process (hence the numbered list above) that builds advanced motor skills by adding component by component, using feedback to shape and improve the execution until a smooth, automated action is reached.

1.7 Learning curves

The process of acquiring surgical skills or a specific surgical procedure can also be expressed by learning curves (80,85). The learning curve of any procedure can be defined as a repetitive practice until the procedure is mastered. A learning curve usually consists of an initial steep phase during which the ability to complete the task increases rapidly. This slope of progress then changes slowly, whereby the improvement in

outcome becomes more modest until the curve becomes flat (plateau) without any further detectable change.

During the first phase of learning a new laparoscopic procedure the risk of errors or experience serious complications is high (86,87). In addition to this, the first part of the learning curve is usually associated with longer operating time and higher conversion rates; both factors also contributing to higher costs. The use of laparoscopic simulators for training basic skills is aimed at shortening the learning curve and improve patient care by providing a safe and controlled environment, thus producing safer surgeons faster (88). In a master-apprentice model proficiency is considered reached when the tutor/supervisor regard the trainee «ready» for performing surgery independently. This variable and highly subjective decision is prone to bias. Estimations in the literature and in guidelines on how many procedures necessary before reaching proficiency varies greatly, and the number is usually relatively arbitrarily chosen, and the level of agreement amongst experts is often low (31). These empirical numbers are often based on expert opinion, and not underpinned by scientific research. In a systematic review by Dagash et al. it was demonstrated that a variation between 8 and 200 for laparoscopic cholecystectomy, and between 20 to 60 laparoscopic fundoplications was considered necessary before reaching proficiency levels (89). The study compared several common laparoscopic procedures and concluded that there was no agreement on how many procedures a surgeon needs to perform before reaching proficiency.

Simulator-based training has, however, provided the opportunity to objectively assess performance, generate learning curves and subsequently evaluate proficiency based on

previously set performance levels (80,85,90,91). Learning curves are individual (92,93) and procedure specific (94,95), and reliable methods or tools for prediction of the future slope is currently lacking. Learning curves are usually created after the training has been completed, and will therefore have no practical value during training of the individual.

1.8 Objective assessment of surgical competence

In 1978, Spencer postulated that the manual skills of the surgeon represented 25% of a competently performed procedure, the remaining 75% representing decision-making (96). In laparoscopy, the dexterity skills probably account for a higher percentage. To be able to evaluate the effect and impact of surgical training, tools that can adequately, reliably and feasibly measure surgical performance are required. A multifactorial variable such as surgical competency is, however, challenging to measure. Ideally it would be assessed with increasingly precision if attacked from several viewpoints (97-99). In numerous attempts to define and measure this elusive variable, several methods and concepts have been developed for use in assessment and training of surgical technique. These objective assessments contain both quantitative and qualitative components.

Since feedback on performance seems to be crucial in training and improvement of skills, objective assessment of operative skills is essential to obtain (76,100). Reliable assessment variables are also necessary to make up the basis for certification processes or other high-stakes evaluations (79,101).

The traditional direct observation of performance in the OR, as a method of assessing surgical technical skills, is highly subjective. This method represents a global evaluation, not based on objective criteria and is therefore coloured by, and dependent of the observer's present relation towards the trainee. Previous studies have shown poor test-retest and inter-observer reliability (even amongst experienced senior surgeons), which leaves this method unsuitable and unreliable for laparoscopic training (102). When laparoscopic cholecystectomy first was introduced, proctoring by experienced laparoscopic surgeons was advocated by the international surgical societies before granting operating privileges. Adding objective criteria by way of a carefully designed and validated checklist, examiners acted as observers, rather than interpreters of behaviour, thus eliminating the subjective from the evaluation. Consequently, objective feedback ensured successful proctoring before the surgeons commenced unsupervised surgeries.

1.8.1 Checklists and global scores

One of the earlier, and probably the most extensively used, assessment of surgical technical skills is OSATS, «Objective Structured Assessment of Technical Skills», established by a group from Toronto, Canada, led by prof. Reznick. This assessment consists of six stations where surgical trainees perform different segments (or part-tasks) of procedures on live animal tissue and bench-top trainers within a predefined time frame (103,104). An expert surgeon present at each station evaluated the trainees by using task-specific checklists and global rating scales, with a high interstation reliability and construct validity.

Global ratings scale (GRS) is a list of seven universal components of surgical skills, denoted from 1 to 5 on a Likert scale, the highest score signifies best performance. The middle and end-points on the scale are explicitly defined (Figure 5) aiding the assessors by providing clear criteria for the assessment (104). Procedure specific GRS's are also developed to evaluate specific technical aspects for procedures as laparoscopic cholecystectomy (105) and Nissen's fundoplication (106). GRS has shown to be a valid tool for evaluating surgical technical skills (107).

GLOBAL RATING SCALE OF OPERATIVE PERFORMANCE

Please circle the number corresponding to the candidate's performance in each category, irrespective of training level

Respect for Tissue :				
1	2	3	4	5
Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissue appropriately with minimal damage
Time and Motion :				
1	2	3	4	5
Many unnecessary moves		Efficient time/motion but some unnecessary moves		Clear economy of movement and maximum efficiency
Instrument Handling :				
1	2	3	4	5
Repeatedly makes tentative or awkward moves with instruments by inappropriate use of instruments		Competent use of instruments but occasionally appeared stiff or awkward		Fluid moves with instruments and no awkwardness
Knowledge of Instruments :				
1	2	3	4	5
Frequently asked for wrong instrument or used inappropriate instrument		Knew names of most instruments and used appropriate instrument		Obviously familiar with the instruments and their names
Flow of Operation :				
1	2	3	4	5
Frequently stopped operating and seemed unsure of next move		Demonstrated some forward planning with reasonable progression of procedure		Obviously planned course of operation with effortless flow from one move to the next
Use of Assistants :				
1	2	3	4	5
Consistently placed assistants poorly or failed to use assistants		Appropriate use of assistants most of time		Strategically used assistants to the best advantage at all time
Knowledge of Specific Procedure :				
1	2	3	4	5
Deficient knowledge. Needed specific instruction at most steps		Knew all important steps of operation		Demonstrated familiarity with all aspects of operation

Figure 5 - Global Ratings Scale of Operative Performance

OSATS, checklists and GRS represent a feasible, but time-consuming, expert-demanding method (103), even if the assessors use time edited video recordings to eliminate the need for multiple faculty members to be present at a specific time and place. However, video-assessment adds further objectivity by blinding the assessment process, and are thus preferred and recommended. A study by Dath et al. et al. also demonstrated that video-assessment time could be reduced by up to 80% by letting expert laparoscopists fast-forward through the unedited videotapes of surgical trainees performing laparoscopic Nissen's fundoplication and a low anterior resection of the rectum in a pig (106). The assessors viewed and fast-forwarded the tapes at their own discretion, whilst scoring by GRS and an OCRS-checklist (Objective Component Ratings Scale). This method showed reasonable inter-rater-reliability (106). OSATS and GRS are also beneficial in providing an opportunity to identify errors and serve as valuable objective feedback to the trainee to decrease deficiencies in training and performance.

1.8.2 Dexterity analysis

Dexterity analysis (or motion analysis) has evolved from motion psychology where the assumption (and empirical observation) was that an experienced laparoscopic surgeon would use fewer and more accurate movements to perform a task compared to an inexperienced surgeon (108). Dexterity analysis, either by tracking instruments or hand movements, is technically possible when training on a simulator, box trainer or during a real operation in an OR. Assessment tools can be built-in (e.g. ProMIST™ or LapSim™) or used as an add-on function (ICSAD/ROVIMAS). Given the magnitude of available simulators and box-trainers, several sets of motion analysis parameters (MAP) have

been developed to describe and assess different aspects of motions performed with laparoscopic instruments. This creates difficulties comparing results from performances across different laparoscopic simulators, and the variables measured are often defined and calculated in a unique fashion for each simulator. The lack of valid, reliable and uniform MAPs demands on-going research into the interpretation of each motion variable and the corresponding predictive value.

A recent review study by Mason et al. claims that there is sufficient evidence that motion analysis can be used for laparoscopic skills assessment, but there is still a lack of predictive validity studies (109). This study also concludes that the MAPs «time taken», «path length» and «number of hand movements» are the most valid in indicating higher levels of performance («better surgical performance»).

1.8.3 Analysis of the final product or outcome after surgery

Surgical skills can also be assessed by evaluating the final product. This can be exemplified by clinically relevant measurements like intraluminal diameter or leak-point pressure of a sutured anastomosis, or the mechanical stability and strength of a surgical knot made by a trainee (110,111). These assessment methods usually require specific equipment and are, therefore, unpractical and unfeasible in a clinical setting, as they are only achievable within the settings of a research project or as a part of a course for laparoscopy training in specific simulation laboratories. Analysis of surgical outcome to evaluate technical skills raises specific problems and it is often difficult to clearly attribute to poor surgical technique. Adverse events after surgery may also go undiscovered until several years after the procedure, and there are too many factors

involved, making it impossible to deduct the impact of surgical technique. One way of overcoming these limitations are performing assessments by using only standardised tasks, for example bench top models (112,113). This approach has demonstrated construct validity and significant correlation between outcome analysis and assessment by OSATS (113).

As mentioned previously, the «perfect» objective assessment method, or set of MAPs, is not established, and there are multiple reasons for this. In 2001, the Metrics for Objective Assessment of Surgical Skills Workshop gathered an international, multi-society board of surgical educators and researchers together with representatives for different official bodies responsible for surgical education, evaluation and certification (114). The workshop was instituted as an attempt to establish a nomenclature and standardize the assessment methods necessary to form a common ground and taxonomy for the surgical educational communities. In addition to this, the workshop was aimed at providing a model for the core curriculum in laparoscopic training. The specific goal of the workshop was to use all current available research to develop a consensus for a baseline set of metrics. These metrics were to be used for future training, assessment and research concerning all aspects of surgical technical skills (114). The workshop resulted in a list defining expressions commonly used by the community, a comprehensive taxonomy (classification) of all the tasks, skills, abilities and procedures involved. In addition, the workshop provided a list of the available (validated) training systems and the corresponding abilities which they offered training for. The workshop did not succeed in producing a core curriculum model. In the final report the participants admitted that this goal was «a bit overambitious», given the available time

and on-going discussions about what the curriculum was intended to be used for, and by whom. However, they agreed on a list of skills, tasks and procedures and succeeded to categorise each item on the list into three levels; basic, intermediate and advanced to provide programme directors with a framework and aid in programme development and training. Finally, several specific research areas were identified as well as an intention to organise a future open forum to maintain the focus and advocate national and international participation in this field.

Great innovations in medical technology lead to a constant advancement of surgical technique. The definition of surgical proficiency will consequently change and development of methods of assessing surgical skills must follow in the same direction. Nonetheless, how we measure surgical skills will always be dependent on, and influenced by, the present available assessment methods. The workshop was considered an initial approach and future meetings were suggested to continue refining and developing a consensus in this area. As of today, a follow-up conference has not been organised.

1.9 Validation of surgical simulators

Before simulators can be applied in a surgical educational system, they must be scientifically evaluated, rigorously and objectively, to determine their validity and reliability. Validity, in the context of surgical simulators used as tools for training and assessment, describes to what extent the simulator meets its requirements and whether it fulfils its intended purpose (115). Given the complexity of the validation concept and procedure, there are several types of validity involved; however, in the validation of

surgical simulators the most commonly used validation terms are face validity, content validity, construct validity, concurrent validity and predictive validity (115). Face and content validities express the translational validity (or representational validity); to what extent the measures can capture and turn an abstract theoretical construct into a specific, practical measurement or test.

- **Face validation** is an estimation of how strong the tested instruments or models resemble the real-life situation that is supposed to be emulated. Face validity does only relate to whether the model/test *appears* to be a good measure, and offers no guarantee that the estimation is correct. Thus, face validity is not a strong validation method as it is based on experts' opinions and subjective emotions. Testing for face validation usually involves administering a questionnaire to experts of the field in question. In this thesis, face validation will be used to test whether training and assessment on a box-trainer called D-Box correctly imitates/resembles training and assessment in the OR. Because of its subjective nature, it is usually only used in the initial phase of the validation process.
- **Content validation** refers to relevance of the modality used. Content validation tests how well the content of the measurement represents the content of the domain being tested, if it contains all relevant aspects. This requires a good knowledge of the field of interest and it usually involves a panel of experts being asked to review each item included in a test (or

assessment tool), and is usually performed early in the validation process, e.g., as a part of a pilot study.

- **Construct validation** refers to whether a test can measure the abstract construct that underlies the qualities it was designed to measure.

Construct validation can be defined as «*a set of procedures for evaluating a testing instrument based on the degree to which the test seems to identify the quality, ability or trait it was designed to measure*»

(115). When used in validation of surgical simulators this is usually performed by measuring the performance of different groups that theoretically should differ in the skills being measured (e.g experienced surgeons vs. medical students performing surgical tasks). This makes it one of the most important, and mandatory, validations for surgical simulators.

In addition to translational validity, which contains a subjective evaluation of the assessment tool's (or test's) resemblance to the physical situation it is supposed to evaluate, and its appropriateness in the intended use, there is a need for objective validation methods. Criterion validity describes to which extent the acquired measures correlate to other measures or outcomes. Criterion validity can be subdivided into concurrent validity and predictive validity:

- **Concurrent validity** is often reported when developing a new test or assessment tool, and the validation refers to how well the new test correlates/compares to the established «gold standard».

- **Predictive validity** can be defined as «*the extent to which the scores on a test are predictive of actual performance*» (115) and relates to whether a measurement (or test) at one point in time will predict (future) outcome or result. When this term is used in the context of surgical simulators, it usually denotes if the level of performance on the simulator corresponds to the level of performance in the OR. That is why it is perceived as the ultimate, or highest-level of validation for surgical simulators. Predictive validation is regarded as the validation method with the greatest clinical impact and relevance.

1.10 Reliability

Validity is commonly evaluated together with reliability. Reliability (r) is a broad and generic expression that usually incorporates expressing the level of consistency of repeated measurements from an assessment (tool) under comparable conditions.

Validity and reliability are equally important. However, they describe different aspects, or properties, of the tests. There are several methods (classes of reliability) developed to establish reliability, the most relevant to this thesis are:

- **Test-retest reliability** assesses to which degree test results are consistent between different repetitions of the same test, administered to the same subject/rater, under the same conditions, but at different occasions/times.
- **Inter-rater reliability** describes the degree of agreement between two or more raters, rating the same subject or evaluating the same test.

- **Inter-test reliability** describes to which degree test scores are consistent across a variation in methods or instruments used.
- **Internal consistency** is often used in the literature of surgical simulation describing consistency across different variables within a test. The most commonly used measure for internal consistency is Cronbach's α . It can easily be calculated for any data set using most available statistical software packages, like IBM SPSS Statistics® (IBM, New York, USA).

All mentioned classes/methods generate a reliability coefficient (r), ranging from 0 to 1, where a value of $r > 0.8$ usually is accepted as good reliability.

It is important to acknowledge that an assessment tool with a high reliability coefficient, does not guarantee a «good» test used in any situation, but one must always evaluate what the test really measures and in what setting (validation). Another important point to realise is that reliability and correlation are not the same, and that correlation simply denotes association (some degree of relationship between the measurements) and reliability is a measure of agreement (equal measurements, likeness). Reliability does not imply validity, however a lack of reliability limits the overall validity of a test.

1.11 Laparoscopic basic skills curricula

There has been a vast interest in validation of newly developed simulators in the last two decades, and a substantial amount of work has been done in this area (46,61,90,116-118). However, the effectiveness of any simulator-based training programme depends mostly on the applied curriculum, and not on the type of simulator

used (119). Several initiatives have been launched to create curricula for teaching and assessing laparoscopic skills, however, no consensus has been established, neither on a national level nor internationally, on what the most efficient training curriculum would include, and what part simulation training would play in such a curriculum. In the USA, the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) have developed a curriculum entitled «Fundamentals of laparoscopic Surgery» (FLS) (79,120). This curriculum is endorsed by the American College of Surgery and is one of the requirements for certification by the American Board of Surgery (18,121). The curriculum consists of CD-ROM-based learning modules (cognitive skills) together with a standardised manual skills training and assessment component consisting of five tasks performed on a custom-built box-trainer similar to the D-Box (79). The FLS performance metrics (time and precision) are comprehensively validated (79,122) and studies show that performance on the simulator improves clinical laparoscopic skills (101), and the performance correlates with performance in the OR (123). This is probably the most widespread basic skills laparoscopic curriculum in North America, however, certification is voluntary and about 27.000 surgeons (of approximately 136.000 members of American College of Surgeons in 2009) currently hold time-limited certifications (121,124). There are similar initiatives originating from Europe as well. In Sweden, simulation courses aiming at certifying surgeons are being organised (CLK), and The European Association for Endoscopic Surgery (EAES) have developed, validated and are in the process of implementing a training curriculum (LSS) for selected laparoscopic procedures along with an E-learning platform (125). Still, no nationally widespread certifying curriculum, including skills assessment, is established in any country (126,127).

None of the previously mentioned basic skills curricula has gained any significant market share in Norway, however, the D-box is present at many teaching hospitals in Norway, thus there was a demand for an evidence-based curriculum and expert-derived proficiency-levels for this simulator.

2.0 General aim of the study

The overall aim of this PhD study was to design a feasible, evidence-based curriculum for training basic laparoscopic skills using the D-Box – a trainer box simulator that currently is available in most Norwegian hospitals and surgical training centres.

2.1 Specific aims

1. To assess volume-outcome relationship in the light of quality of care and surgical education.
2. To validate D-Box for its usability for training and assessment of basic skills in laparoscopy.
3. To assess transferability of skills between D-Box and MIST-VR, representing a trainer box and a VR-trainer.
4. To create an evidence-based curriculum for the validated, available tasks, using scientific methods.

3.0 Materials and methods

For Paper II-IV the hardware, training and assessment setup for the D-Box were identical. These studies were executed at Ostfold Hospital Trust, at both hospital locations: Moss and Fredrikstad. Paper II also included use of MIST-VR. All training and assessment sessions were carried out in a quiet room, under standard ergonomic settings and with the same instructor/supervisor present for all sessions. Training sessions were organised so that the trainees attended sessions whenever they had available time during their clinical work, to resemble training opportunities in a normal clinical setting.

3.1 Simulators used in the studies

3.1.1 MIST-VR®

The first commercially available laparoscopy VR-simulator was the MIST-VR® (Mentice AB; Gothenburg, Sweden, www.mentice.com) (128). At the time of commencing these studies the MIST-VR® was also the most extensively validated VR-simulator and incorporated into surgical basic skill courses internationally (56,57,86,91,92,128-135). For paper II we utilised MIST-VR®.

Figure 6 - MIST-VR[®] Setup and display

The user interface (Immersion Inc.; San José, USA) was linked to a computer and had a frame that held two specially designed mock-up laparoscopic instruments with interchangeable handles. Both instruments pass through a position-sensing device providing real-time positional data in 5 degrees of freedom. Positional data were then processed by the computer and bespoke software, which generated a real-time graphical image of the virtual instruments and their movements presented on a 19" colour monitor. The displayed image was a three-dimensional cube as the abstract operative space, measuring exactly 10 cm^3 . During training different targets appeared randomly within the cube, offering distinctive challenges to the surgical trainee. The software provides the opportunity to zoom in/out, and to change the size of the targets on the screen. In the setup used during this PhD research, the MIST-VR[®] system did not include force feedback or haptics.

The MIST-VR[®] software was delivered with two preinstalled basic skills modules (Core skills I and II) and a suturing module (Suture 3.0). Only Core Skills I was utilised in our studies. Core Skills I contained six different tasks (Table 2), and each task can be set to three levels of difficulty; Easy, Medium and Hard. The MIST-VR[®] tasks are previously validated by Gallagher et al. (131).

No.	Task name	Description	Operative task
1	Acquire-place	Grab a sphere and place it within a wire-frame cube.	Manipulating tissue and bringing it to a given position.
2	Transfer-place	Grab a sphere with one hand, pass it to the other hand, and place it within a wire-frame cube.	Manipulation of a needle during intracorporeal suturing.
3	Traversal	Use instruments to travel from top to bottom along the outside of a three-dimensional cylinder.	Running through the small intestine.
4	Withdraw-insert	Pull an instrument out of the operative field and reinsert it for further use.	Withdrawing and reinserting new instruments into the abdomen.
5	Diathermy	Cauterize a series of targets located on a sphere fixed in space.	The use of diathermia on a bleeding vessel.
6	Manipulation-diathermia	Grab a sphere and touch it with the instrument in the opposite hand; withdraw and reinsert the opposing instrument; apply diathermy to targets on the sphere while holding the sphere steady within a wire-frame cube.	Using diathermia to dissect the gallbladder off a liver bed during a laparoscopic cholecystectomy.

Table 1 - Core Skills I (MIST-VR[®]). Description of each task and correlated operative task.

During training, the performance of the trainee is monitored and recorded in real-time. The collected data includes movements of instruments placed in each hand and the use of a foot pedal (during the more complex tasks) to represent the use of a foot-activated electrocautery device. After each training session the MIST-VR[®] software provides the trainee with detailed feedback of performance, and a video playback of the session can be reviewed by the trainee alone or together with a supervisor. The automated feedback metrics included time to complete the tasks (TTC, in seconds), Economy of movement (EOM, in numbers) and total score (TSC, in points) (86). All metrics were given for

individual hands, but left and right hand performance were added into total values before analysis for uniformity.

3.1.2 D-Box

The D-box trainer (SimSurgery, Oslo, Norway, www.simsurgery.com) has been continually developed since 2003. The model used in our studies were the second generation design and was constructed of an aluminium box measuring 32 cm x 45 cm x 21 cm (width, depth, length) and weighing 10.5 kilograms. A slideable tray in front of the trainer box enables insertion of different task boxes. During these studies, five different task boxes were used for a total of six tasks.

Inside the D-Box there are 14 LED-lights and a web camera mounted in the roof. The web camera allows external manipulation by a joystick, omitting the need for an assistant controlling the camera. The web camera connects to a laptop computer via an USB-connector and the video image was presented on a 17» monitor.

Two commercially available graspers (EndoPath®; Ethicon Endo-Surgery, www.ethicon.com) were used for all tasks in the D-Box.

Figure 7 - D-Box: Different task boxes available

No.	Task name	Description	Operative task
1	Peg Transfer	Sorting eight coloured pegs into two different boxes, alternating left and right instrument.	Basic manipulation of tissue, using one instrument at a time.
2	Sorting Pegs	Pick up and pass the eight coloured pegs through an eyebolt before placing the pegs into to boxes, according to colour.	Basic manipulation of tissue, using both instruments.
3	Donkey Stack	Using both instruments, stacking five wooden sticks simultaneously on the back of a figure of a donkey.	Precision manipulation of tissue using both instruments simultaneously.
4	Running gut	By using both instruments to run through 170 cm of silk ribbon, from one corner of the box to another.	Running through the small intestine.
5	Rubber plate	Picking up and passing a pin through five predefined holes in a vertically mounted, movable rubber plate, alternating from left to right.	Manipulating tissue and bringing it to a given position.
6	Labyrinth	Passing a peg through seven eyebolts in prefixed directions and angles, at one point lift a rubber plate and passing through a hidden eyebolt, before finally threading the pin through the ninth eyebolt in an opposing direction.	Manipulation of a needle during intracorporeal suturing.

Table 2 - D-Box task boxes. Description of each task and the correlated operative task.

3.2 ICSAD/ROVIMAS – *electromagnetic tracking of hand motion*

As opposed to VR-simulators, box-trainers are usually delivered without any other objective measurement options other than timing how fast the trainees can complete each task. A number of tools for objectively assess technical performance have been developed, but for automatic measurements of surgical dexterity different methods of tracking hand- or instrument movements are suitable (136). Thumb-sized electromagnetic coils (receivers) were attached to a tight-fitting, but comfortable, cotton glove to a standard position on the dorsum of each hand of the trainees. A electromagnetic emitting device, also a part of the commercially available electromagnetic tracking device, Isotrak II (Polhemus Inc; Colchester, USA, www.polhemus.com), then provide real-time Cartesian positional data (X, Y and Z coordinates), thus generating three-dimensional (3D) coordinates for any object's position in space. These 3D-coordinates, or rather their relative movement, were recorded at 60 Hz (divided between the two receivers) and transmitted to a laptop computer, imported into the Imperial College Surgical Assessment Device (ICSAD) to add time-stamps and translate the Cartesian data to the objective variables of individual hand movements. We utilised the previously validated parameters; time taken to complete task, number of hand movements, total path-length (62,137,138). Parameters were specified for each individual hand, but left and right hand parameters were added together before analysis. These measurements were later analysed by ROVIMAS (Robotic Video Motion Analysis Software), a custom-made software-package also developed at Imperial College, London, UK (see Dosis et al. for details (139)). The software uses built-in noise filters to detect and filter out tremor and unintentional movements. Later versions of the software also included the possibility to synchronize

and combine video recording (and thus evaluation) of the procedure (or just specific parts of it) with motion analysis for improved assessment and feedback (139). This also allows «zooming in» on key areas of the procedure for faster video assessment combined with dexterity analysis data for that specific part of the procedure, for example clipping and cutting of the cystic duct.

The assessment method using ICSAD/ROVIMAS is extensively validated and is used for assessment of dexterity in laparoscopic, microscopic and open surgery, basic skills and suturing, and in different surgical fields including ophthalmic, orthopaedic, cardiac and plastic surgeries (97,112,140-144). The assessment method has also proven feasible for the OR-setting (97).

Figure 8 - Setup for ICSAD electromagnetic coils

3.3 Data access and ethical considerations

All participants in paper II-IV were recruited after receiving a letter of invitation containing information about the project and then actively contacting the study administrator to volunteer for participation. A written consent form was obtained from all participants.

The local Regional Committee for Medical Research Ethics (REK) were contacted regarding all our studies separately, concluding each time that our studies fell outside their regulatory mandate, and a formal, written application was thus never required or submitted.

All data were stored and processed after application and subsequent approval (Ref. no. 15254/2006) from Norwegian Social Science Data Services (NSD) and in accordance with guidelines provided by the Norwegian Data Inspectorate (Datatilsynet).

3.4 Statistics

Statistical analysis was performed using SPSS Statistical Package, version 14.0 and 17.0 (IBM SPSS Statistics®, IBM, New York, USA). Differences between proportions in contingency tables and frequency tabulations were analysed using Pearson's X^2 (where appropriate Fisher's exact test was applied). Differences between normally distributed variables were analysed using Student's t-test. All numeric variables were considered non-parametric and all probabilities were calculated as two-sided and the level of significance was set at $p \leq 0.05$.

- **Paper I:** For 5-year cumulative survival rates the Kaplan-Meyer method with log rank probabilities for survival comparison between groups was applied. The survival rates are given as crude survival.
- **Paper II:** Mann-Whitney U test were used for comparison of performance between the two groups. Kruskal-Wallis test was used for evaluating the training effect across the training programme and Wilcoxon signed rank test for related variables.
- **Paper III:** Mann-Whitney U test were used for comparison between the three groups.
- **Paper IV:** For learning curve evaluation the Friedman test (non-parametric, repeated measures for analysis of variance) were applied. Multiple comparisons were then made to identify the point of plateau (no further improvement).

4.0 Synopsis of papers and individual main conclusions

4.1 Paper I

«Curative rectal cancer surgery in a low-volume hospital: A quality assessment.»

Objective:

To assess the relationship between volume and outcome for rectal cancer patients in a low-volume hospital, illustrated by a 10-year retrospective material from our hospital, compared to results from the National Rectal Cancer Registry (NRCR). This study was initiated as a retrospective internal quality review to assess oncologic outcome and the impact of individual surgeon procedure volume.

Participants:

By using relevant diagnose codes (ICD-9: D153.3 and D154, and ICD-10: C19 and C20) we performed a search of our hospital electronic database. All patients diagnosed with carcinoma of the distal 15 cm of the rectum in the period 01.01.1993 to 31.12.2002 were identified. All patients having surgical treatment (both elective and emergency operations), with curative intent, were included. Their records were manually explored and a total of 131 patients were selected. Patients having endoscopic resections were excluded, as well as patients receiving primary palliative treatments. Finally, we compared our patient list with a list provided by the NRCR, bilaterally correcting any discovered discrepancies. The final study population consisted of 102 patients. Mean patient time at risk was 58 months.

Method:

We performed a retrospective review of relevant, available medical charts for each patient. The list of patients was compared to the list from the National Cancer Registry for completion of records bilaterally.

Results:

There were no significant differences from the national outcome results, neither in perioperative mortality or complications, nor 5-year survival or local recurrences. Thirteen different on-staff surgeons performed rectal cancer surgery in our hospital during this decade. Median annual caseload as the main surgeon was four. We detected a difference in 5-year survival when grouping the surgeons by annual caseload, but the significance is inconclusive. It is, however, interesting that in 85% of the resections, two or more certified gastrointestinal surgeons with specific training were involved. The study revealed a 9% discrepancy between records from the Norwegian Rectal Cancer Registry (NRCR) database and the local hospital database.

Conclusion:

Adequate results for surgical outcome can be achieved in a low-volume hospital. Surgeon volume showed inconclusive impact for our results of outcome. A local quality initiative is justified in addition to national registries.

4.2 Paper II

«A tale of two trainers: virtual reality versus a video trainer for acquisition of basic laparoscopic skills.»

Objective:

There is increasing evidence that training with laparoscopic simulators is efficient in developing skills that improve performance in real operations. However, the best modality for training novice laparoscopic surgeons is not yet established, and studies comparing box-trainers and VR-simulators show contradictory results. The aim of this study was to assess the transferability of skills between the two modalities, represented by D-Box and MIST-VR[®].

Participants:

38 medical students with no experience in laparoscopy were included and electronically randomized into two groups, A and B.

Study design

Group A received 8 sessions of training on 5 tasks with pre- and post-assessment on the MIST-VR, with a final crossover assessment on the D-Box. Group B received 8 sessions of training on 5 tasks and pre- and post-assessment on the D-Box finishing with a final crossover assessment on the MIST-VR[®]. For performance on the MIST-VR[®] the built-in metrics were used, including time to completed task (TTC), economy of movement (EOM) and Score. Performance on the D-Box included time to completed task (TTC), total number of hand movements (THM) and total path length (TPL), and

was recorded using the Isotrak II ® electromagnetic tracking system and analysed using validated software (ROVIMAS/ICSAD).

Results:

There were no baseline differences between the groups in age, sex, handedness or previous experience with simulators. Both groups significantly improved their performance assessed before and after training for all variables. When tested on MIST-VR®, the MIST-VR® group showed significantly shorter time (90.3 vs. 188.6 seconds, $p<0.001$), better economy of movements (4.40 vs. 7.50, $p<0.001$) and lower score (224.7 vs. 527.0, $p<0.001$). However, when assessed on the D-Box there was no difference between the groups for time (402.0 vs. 325.6 seconds, $p>0.15$), total hand movements (THC) (289 vs. 262, $p>0.79$) or total path length (TPL) (34.9 vs. 34.6 meters, $p>0.38$).

Conclusion:

Both simulators provide significant improvement in performance. Our results indicate that skills learned on the MIST-VR® are transferable to the D-Box, but the opposite could not be demonstrated.

4. 3 Paper III

«Construct, content and face validity of the D-Box; a web-cam based laparoscopic basic skills trainer box.»

Objective:

Basic laparoscopic skills can be acquired feasibly and safely using simulators. The D-Box trainer is a box-trainer already available in most Norwegian teaching hospitals and surgical training centres. Prior to implementing D-Box into a training programme it must pass a scientifically based validation process.

Methods:

Study participants were recruited from 3 levels of experience based on their professional title (medical students/interns, residents and consultants). The consultants were all experienced laparoscopic surgeons and were considered experts. 18 interns, 10 surgical residents (PGY1-4) and 10 experienced laparoscopic surgeons were tested on D-Box using 6 bespoke tasks. Performances were assessed using electromagnetic tracking of hand movements (path length and number of hand movements) and time-to-complete task. All participants scored a seven-statement questionnaire on a five-point Likert scale for user perception after assessment.

Results:

97% of all participants would recommend including D-Box as part of a laparoscopic training programme in Norway. The participants were satisfied with D-Box and all user perception statements received a median score of 5, except for screen resolution

(median score of 4). For task 4, TTC showed significant differences between the groups, in favour of the experts, this was also true for all measurements regarding tasks 5 and 6. For tasks 2 and 3, there were no differences between the groups. Task 2 revealed a significantly shorter total path length in favour of the interns.

Conclusions:

This study has established face and content validation for D-Box. In addition, it has been demonstrated that D-Box can distinguish between experienced and inexperienced laparoscopic surgeons for three of the six tasks, and thus establishing construct validity. D-Box did however failed to discriminate between different levels of inexperienced surgeons.

4.4 Paper IV

«Construction of an evidence-based, graduated training curriculum for D-Box, a web-cam based laparoscopic basic skills trainer box.»

Objective:

Increasing numbers of surgical training programmes are now including simulators as training tools for teaching laparoscopic surgery. The aim of this study was to develop a standardized, graduated and evidence based curriculum for the newly developed D-Box for training basic laparoscopic skills.

Methods:

18 interns with no laparoscopic experience completed a training programme on the D-Box consisting of 8 sessions of 5 tasks with assessment on a sixth task. Performance was measured by use of three-dimensional electromagnetic tracking of hand movements, path length and time taken. Ten experienced surgeons (>100 laparoscopic surgeries, median 250) were recruited for establishing benchmark criterions.

Results:

Significant learning curves were obtained for all construct valid parameters for tasks 4 ($p<0.005$) and 5 ($p<0.005$), and reached plateau levels between the fifth and sixth session. Within the 8 sessions of this study between 50-89 % of the interns reached benchmark criterions for tasks 4 and 5.

Conclusion(s):

Based on improvements in time taken to complete tasks (TTC) and hand movements (THM and TPL), benchmark criteria, an evidence-based simulation curriculum has been developed for the D-Box. This curriculum is aimed at training and assessing surgical novices in basic laparoscopic skills.

Figure 9 – Proposed curriculum for training basic laparoscopic skills on the D-Box

5.0 Discussion

Our studies have led to validation of the D-Box and development of a feasible, scientifically based curriculum for training and assessment of basic laparoscopic skills on this simulator. Although not a comprehensive or exhaustive curriculum, and only aimed at training basic skills, it provides a tool for laparoscopic novices to acquire a set of laparoscopic skills before entering the OR for the first time, or assisting an operation.

Training to become an outstanding surgeon can be compared to the long preparations an elite athlete goes through to attain the skills and abilities to win an Olympic medal. Both require long hours of repeated training, insight and knowledge in the field, along with the right attitude or mental state. However, during the training period they both will need the right tools and training environment to develop the necessary skills safely, correct and in the shortest amount of time.

The methodology used in the papers included in this thesis are dissimilar and will therefore be discussed separately, however, the findings and implications will be discussed as a unit.

5.1 Methodological considerations

5.1.1 Paper I

The study described in Paper I was executed as a retrospective review of the available medical charts for each patient who underwent curative resections for rectal cancer in a 10-year period (1993-2002) at one hospital. To acquire the most complete list of

patients to be included in the study, at the authors request the National Cancer Registry provided a comprehensive patient list from their database. These two lists were then compared and information was bilaterally corrected for completeness. The unique individual identification number of each citizen of Norway, combined with compulsory reporting of all cancer diagnoses (from clinical departments and pathology departments), provides a solid foundation for a complete database and enables long-term follow-up. This ensured us that we had included all eligible patients in our study. The limitations of the study relates to the lack of information on neo-adjuvant and adjuvant therapy to complete the assessment of overall performance and recurrences/survival, and the retrospective methodology.

5.1.2 Paper II

The study reported in Paper II was conducted as a prospective, randomised crossover trial. A total of 46 medical/surgical interns and medical students were initially recruited and by computer randomised into two groups. The groups were assigned training on two different laparoscopic simulators receiving a crossover assessment on the other simulator after finishing the training programme. In the paper all subjects are referred to as «medical students» for simplicity and to indicate the laparoscopic naivety. The crossover design was chosen to compare the training effect of two different basic skills laparoscopic simulators. The training tasks included in each simulator were inherently different, but contained all the basic skills; grasping, pick and place, traversing, bimanual instrument handling etc. (114). The training programme on each simulator was designed to be as equivalent in training volume and level of difficulty as possible with the available tasks. The two simulators also differed in their graphic visualisation

and that the trainer-box was the only that provided haptic feedback. It was deemed unfeasible to redesign the tasks in the box-trainer or the VR-trainer to have identical tasks, and it did not represent the choices surgical training programme directors face when investing in surgical simulators.

One limitation with the study design is that there is no baseline assessment of the trainees on both simulators; only from the simulator they were assigned to train on. By adding a baseline assessment for both simulators, the added training effect would be difficult to estimate and account for. However, by employing the crossover design, each subject becomes his/her own control.

5.1.3 Paper III

For this validation paper, based on their professional titles (medical student/intern, resident or consultant), we recruited a total of 40 subjects. All subjects performed a sequence of all six tasks twice, the first for familiarization, and the second for assessment of performance.

Classifying the subjects only by their professional title was a choice of convenience, however, it also correlated with their self-reported levels of laparoscopic experience. A consensus on how to objectively classify surgeons according to their skills does not exist, however there are statistical methods that could have been applied (145). In our results there were difficulties discriminating between the two most inexperienced groups; interns and residents, and their level of experience might be too similar. When

regrouping the subjects into only two groups; the inexperienced (interns and residents) and consultants, the difference in performance increased and the consultants outperformed the inexperienced in almost all tasks. This difference was however not consistently statistical significant and a larger study population could possibly clarify this.

Another limitation with our study design regards the face validity. The questionnaire statements (Q1-Q7) were aimed at assessing the subject's appraisal of the D-Box by marking their agreement to each statement on a five-point Likert scale. The statements covered aspects of technical matter (screen resolution and image delay), the subject's emotions toward using the D-Box (challenging and fun to use) and whether the subjects thought the D-Box improved their laparoscopic skills and if it should be part of training programmes for residents. In hindsight, the statements could also have addressed how realistic the subjects perceived the simulator and if the haptic feedback from the simulator resembled that of human tissue handling.

5.1.4 Paper IV

For this study we recruited a total of 20 medical/surgical interns with no previous experience with laparoscopy or simulation training for laparoscopy. All participants performed eight sessions of five tasks. Before and after the training programme the subjects were assessed on the sixth task. Ten experienced laparoscopic surgeons performed all tasks twice, and benchmark scores were calculated on the basis of the measurements from the second run.

In our study we chose to let our subjects train for eight sessions, based on previous studies demonstrating plateauing of the learning curves before finishing eight sessions (57,76,92). Shortly after planning and initiating this study, several publications emerged, advocating proficiency-based training over pre-set volume training (hours or number of sessions) (85,93,146). Proficiency is defined as achieved when the learning curve plateaus (89), indicating that no significant further improvement is expected. Training until proficiency takes into account the trainee's present skills and performance, tailoring the programme to each trainee and providing efficient training.

By using previously published methodology for training with VR-simulators, the authors proposed a proficiency-based curriculum for training basic laparoscopic skills on the D-Box (147-149).

5.1.5 D-Box assessment

All sessions (training and assessments) on the D-Box within the included studies were performed in a quiet room, under standardised ergonomic setup. Before commencing the first training session, all subjects received a demonstration of the tasks and rules of the study. Two runs of Task 1 (the presumed easiest task) served as familiarisation with the simulator and ergonomic setup before the first session. The training was scheduled during normal working hours, at the subject's own convenience, with a maximum of two sessions per day, separated by at least one hour, to allow a favourable dispersed training distribution (150,151). Practice was not allowed between the sessions. The same study administrator (the PhD-candidate) was available for all training and assessment sessions, providing standardised instructions and technical assistance, before

and during sessions. These precautions were employed to avoid biases in the setup caused by non-random unintentional factors.

Most VR-simulators track and store all instrument movements during training and assessment on the simulator, and immediately after finishing a task or session, report these results to the trainee. It was a natural choice to try to replicate this when assessing performance on the D-Box. The D-Box is distributed without any built-in assessment method, neither hardware nor software applications. To be able to assess the subject's performances we applied electromagnetic tracking of hand movements and measured the time used to finish each task. Most VR-simulators will in addition identify and report performance errors and incorporate this into a total score (152). Tracking and assessing hand movements alone can lead to faulty conclusions if the trainee completes the task with fewer movements, shorter path length and/or in a shorter time just because he/she «cuts corners» or omits significant parts of the task. To ensure consistency in assessments (increase reliability), the study administrator observed all sessions and guided the subjects to perform all tasks correctly and adhering to the study protocol.

The recognition of errors (or potential errors) remains an essential skill for a surgeon, both during training and especially when performing real surgeries. Error assessment was not included as a variable in any of our studies. The use of recorded video would enhance the evaluation/validation by providing an opportunity to address errors by scoring the subject's performance after sessions.

Video assessment, or direct observational assessment, is time-consuming and relies heavily on experts and is usually not feasible outside research programmes (99). An automated computerised assessment is potentially less labour intensive, and thus to be desired. A recent study by Xeroulis et al. demonstrated a high level of correlation between FLS standard score (blinded video scoring by experts) and ICSAD scoring (automated tracking of hand movements) for FLS (122). Although observational assessment still is fundamental, there is rapidly evolving technology/methodology for automated evaluation of surgical skills (99).

5.2 Findings and implications

5.2.1 Volume = quality of care?

In recent years the Norwegian health care system has gone through great organisational transformations. Since 2001, several local hospitals have been closed down, catchment areas increased and larger units (Hospital trusts) have been established within the four state-owned, self-governing health authorities (153). The arguments from the government's side prior to this reorganisation were many, but improving the quality of care was one of the main reasons, in addition to reducing overall costs and thus increasing the number of patients being treated. In the planning process, hospital caseload was one of the variables employed in the planning process and played an important role in deciding what hospital trust would provide which function. Hospital caseload is used in many situations and represents an attractive surrogate measure of quality of care because it is easily obtainable, and it represents a comprehensible quantity. Direct comparisons between hospitals are straightforward. The start of the

work that now has become this thesis was an internal quality assessment of the surgical treatment provided for rectal cancer patients in our hospital. This initiative was raised due to an on-going process where the surgical departments of the two hospitals in our county were to be merged and there was a discussion on whether it was favourable for both hospitals to continue providing surgical treatment for rectal cancer. In addition to performing an internal quality assessment, we set out to investigate if, and what impact, individual surgeon volume had on the results.

Surgical treatment of rectal cancer is a highly specialised function where the patient's outcome is indeed dependent on the surgical skills of the surgeon(s). Multiple studies have shown an association between outcome of different surgical procedures and hospital caseload (154-157). This has also been demonstrated for rectal cancer (25,158,159), however, the association is probably not universal or robust as there are studies showing no association (26,160,161). Also, when adjusting for surgeon caseload, hospitals with low annual caseload demonstrate outcome in the same range as high-volume hospitals, but high-volume hospitals perform better than very high-volume hospitals (159,162,163). In many of the studies, the effect of hospital volume is also relatively modest. There are no consensus on what numbers represents «high-volume» or «low-volume», and these terms are defined and applied very differently from country to country. In Norway, all rectal cancer patients are treated in public hospitals; there are no private alternatives or specialised centres that exclusively treats this patient group. When comparing Norwegian publications of results after surgical rectal cancer treatment to publications from other countries, nearly all hospitals in Norway would be

classified as «low-volume» (25,159,162,163), but still, today both short-term and long-term outcomes are good (163).

This was also the case with the results from our low-volume hospital (as shown in Paper I), where patients experienced adequate and nationally comparable results for 5-year survival rate, local recurrences, perioperative complications and mortality. This in spite of a skewed patient distribution towards more advanced tumour stages and a higher number of patients with lymph node metastasis at time of diagnosis (the reasons for this was not explored in our study). It is commonly recognised that patient outcome after oncologic surgery is largely dependent on the stage and extension of the tumour at diagnosis, in addition to the quality of the surgical procedure performed and the postoperative care (24,164). To maintain an acceptable surgical outcome in low volume units, a special attention towards specialisation, continued training and education is necessary to remediate the lack of operative volume.

Norway was one of the first countries to implement total mesorectal excision (TME) as a national standard for surgical treatment of rectal cancer (165). A national study from Norway had prior to this demonstrated a high frequency of local recurrences following resections for rectal cancer. Surgical technique and individual surgeon performance were identified as possible main contributing factors, due to the wide variations in outcome (166). This, combined with the promising results after TME published by Dr. RJ Heald, led to a national initiative for improving the outcome by initiating a nationwide educational programme. In November 1993, the Norwegian Rectal Cancer Project (NRCP) was founded to improve the outcome (especially reduce local recurrences and

consequently survival) for rectal cancer patients by implementing TME-method as the new gold standard.

To achieve an improvement of outcome after rectal cancer surgery on a national level, regional workshops were organised to provide the training and warrant a standardisation of the operative technique. Dr. RJ Heald, the surgeon that pioneered the TME-method, participated in these workshops as an instructor and supervisor, teaching his method to Norwegian gastrointestinal surgeons (165). The participating surgeons were then certified to perform the TME-technique. In addition to this, pathologists were provided with guidelines on standardised handling and reporting of the removed specimen. A specific rectal cancer registry was also established within the nationwide, obligatory Cancer Registry of Norway (CRN), to monitor the progression and outcome. Later, rectal resections were also moved from the curriculum for surgeons in specialisation for «general surgery» to surgeons sub-specialising in «gastrointestinal surgery». This combined initiative resulted in a 50% lower local recurrence rate from 12% to 6% when comparing TME-surgery and conventional surgery in the study period from late 1993 to mid 1997. The four-year survival rate in the TME-group was 73% compared to 60% in the conventional group (166). These results came about without any organisational changes in annual hospital volume or altered guidelines for adjuvant radio- or chemotherapy. Although many factors could play a role, the structured educational programme, standardisation of the procedure and improvement in the surgical technique was the main «intervention» in this period.

Although we could not demonstrate a consistent surgeon volume effect on outcomes in our study, the majority of the specialised surgeons working in our hospital had participated in the described educational programme. In our study, 85% of the operations were performed by two or more certified gastrointestinal surgeons with specific training in the standardised TME-technique, and this might be one of the main factors contributing to favourable results.

The national rectal cancer registry (NRCR) provides a biannual report for each contributing hospital, where the results/outcomes from that specific hospital are commented and compared to the national average. The average results are also grouped by annual hospital volume. However, each hospital only receives their own results, and comparisons cannot be made to other hospitals. This has just changed, and the August 2013 report from the NRCR is the first to contain a list of outcome results from all Norwegian hospitals providing surgical treatment for rectal cancer².

Although hospital volume is available, at present it is not possible to extract information from the NRCR database to assess annual caseload or outcome for individual surgeons. Initially the surgeon's name and level of specialisation, at the time of the procedure, was entered into the database, but due to poor quality of the reported data, this information was omitted². Some regarded these data as sensitive given the small surgical community in Norway, and rumours were floating around that surgeons deliberately neglected to provide this information out of fear that the information would become public. Recently, the medical director of the National Health Services (NHS) in

² Personal communication: Morten Tandberg Eriksen, Norwegian Rectal Cancer Group

England announced that a league table showing the results of each NHS-employed surgeon will be made public within two years (167). The NHS plans to publish these individual results to «expose variations and unacceptable practice». This initiative came about as an attempt to improve standards by increasing the transparency and thus forces the surgeons to deal with performance and outcome issues actively and openly. Making this information public raise several questions on how they will be applied, and if patients and the public will be able to decipher and interpret the contents, limitations and complexity of these kinds of data? It is well known that only a non-operating surgeon creates no surgical complications, and publishing individual results could lead surgeons to defer from taking on high-risk or complex procedures in fear of negatively influencing their rating statistics. Further, it is not clear yet in which form or forum these NHS league tables are to be published, but mere annual caseload or risk-adjusted mortality rates will in it self only serve as a surrogate measure of the quality of care provided.

Although volume and mortality are easily calculated, measuring the quality of the delivered care, or surgical procedure, is still an elusive task, demanding a range of specific measurements tools. However, validated tools for assessment of competency, even at a specialist level, in laparoscopic colorectal surgery (168) is gradually becoming available and should be further developed and adopted by other specialities. Future assessment and certification of surgeons in training, as well as specialists, must be based on actual and measurable skills and knowledge, not solely on procedural volume.

5.2.2 Validation of D-Box

Tools for training and assessment of surgical skills, to be used for high-stakes decisions, like certification or granting operative privileges, needs to be systematic and thoroughly tested for validation and reliability. The process for validating simulators or curricula for acquisition of surgical skills is advanced, thus no single study will answer all questions of validity (115). A great number of papers describing different levels of validity and reliability of surgical teaching tools have been published in the last few decades, each providing a small piece of the puzzle. Still, the ultimate simulator or curriculum remains unknown.

In our studies we assessed motion economy (total path-length and number of hand movements) and time to complete specific tasks as objective indicators of surgical skills. In the available literature, a myriad of different variables (or metrics) have been developed, and different setup and methodology have been applied in similar studies (109,169). A great number of these validation studies lack specific details on the setup, leaving unanswered questions to what exactly has been validated. A recent study by Våpenstad et al. compared a previously validated proficiency-based training programme consisting of two tasks, using LapSim®. The setup included using haptic feedback handles (170). Previously, construct validity had been established for the identical simulator and identical tasks, using regular, non-haptic handles. The haptic-setup failed to confirm construct validity for 18 of 19 metric variables, indicating that changing details in the setup can lead to totally different results. Further studies may also show that certain metrics are only valid for certain setups, and that not all MAPs are suitable for all setups or tasks. This creates a need for a great number of validation studies to

reveal the validity of each MAP. Based on the available literature, motion efficacy metrics and time measurements demonstrate best validity and reliability when assessing laparoscopic technical skills (99,109,169).

5.2.3 Transferability of skills

A Cochrane review from 2009 summarises the available publications and concludes that there is accumulating evidence supporting VR and box-trainer simulation training as a valuable supplement to traditional laparoscopic training, and that VR-training is at least as effective as box-trainers (171). A 2013 update (172) has addressed the effect of simulation training on patient outcome (benefits and harm) by selecting only trials including subjects with limited laparoscopic experience prior to simulator training. The authors conclude that VR-training appears to reduce operating time and improve operative performance in inexperienced laparoscopic surgeons, however the impact of these effects on costs and patient-related outcome is still unknown. Both studies concludes that further research, of better methodological quality, is necessitated and the impact on clinical outcome needs to be addressed (171,172). In addition, both reviews express difficulties comparing results from different studies given the great variation in measurement methods and reported outcome variables (MAPs). There is still no consensus on what simulator (or simulator type) provides the most efficient and long-term acquisition and retention of laparoscopic skills. Evidence is accumulating that VR-simulators and box-trainers complement each other, however single-modality or multimodality curricula for training basic laparoscopic skills demonstrate similar outcome (66,173).

Prior to commencing our study (Paper II), only a few researcher groups had assessed the transferability of skills between the two basic types of simulators (VR vs. box-trainers) in an attempt to determine their correlation and internal level of efficacy for training basic skills to novice surgeons (174,175). Our study supports previous findings that both types of simulators provide significant improvement in performance across the training period. However, when comparing the post-training crossover assessments, the MIST-VR®-trained group performs significantly better than the D-Box trained group, implying that the skills learned with the D-Box does not match the skills learned with MIST-VR®. The MIST-VR® trained group performed on the same level as the D-Box-trained group when tested on the D-Box. The inherent differences in the two simulator types create several possible explanations for this, including the use of non-identical tasks and difference in haptic feedback. Unfortunately, our study was not proficiency based, and additional training volume on the D-Box could therefore influence the results. Another recent crossover study published by Loukas et al. compared identical basic skills tasks on a VR-simulator and a box-trainer demonstrated transferability of skills both ways (173). The study setup was slightly different, given that the novice participants (n=44) were divided into two groups training on either a VR-simulator, or a box-trainer, receiving pre-training and post-training assessment only on the other simulator. Although both simulators lead to improvement in performance metrics, similar to our results, training on one modality did not necessary lead to an equivalent performance level on the other modality (173,176).

Transferability of skills between different training modalities depends on many factors, difference in haptic feedback has been proposed as one (169,170,177-179). Some

reports conclude that simulators equipped with realistic haptic feedback (VR or physical) should be used for tasks requiring the application of force, like suturing and knot tying. Other tasks, requiring mostly positioning of the objects and instruments, can be trained on simulator systems without such haptic feedback (178).

The ultimate goal and method of testing for transferability of skills (and thus denoting training efficacy), is to compare performance, after training, performing real cases. The transfer of skills from VR (or box-trainers) to the OR has been demonstrated for several modalities (51,60,62), and the choice of which modality to be used for laparoscopic simulation training can be regarded of less relevance, as long as the chosen device is efficient in shortening the learning curve on real cases (180).

5.2.4 D-Box curriculum and benchmark criteria

Our studies have led to the creation of a graded curriculum aimed at training novice surgeons basic laparoscopic skills, using the D-Box. Our proposed curriculum addresses only the technical aspects of basic laparoscopic skills, and is thus just one of the components in a full laparoscopic curriculum. It is important to remember that surgical competency is a multifactorial variable and future curricula must also seek to include skills and knowledge from the fields of theoretical knowledge, communications skills, team leadership and surgical decision-making. Procedural training should naturally follow basic skills training, and several VR-simulators and box-trainers now offer procedural training (46).

Prior to synthesising our curriculum, a set of benchmark criteria was established by use of experienced laparoscopic surgeons. These criteria facilitate proficiency-based training, which takes into account that trainees follow an individual learning curve, due to their inherent abilities and previous experiences (90,146). The use of proficiency-based training ensures that technically skilled trainees finish their training when they are proficient, without prolonging the training unnecessarily. Subsequently, resulting in prolonged training for those not able to meet the criteria within a certain number of sessions. This ensures the most efficient training programme for all trainees, and is thus recommended (176).

When creating training curricula it is important to adhere to scientific principles, and our curriculum was based on methodology previously used in VR-simulator curricula design (147-149), following the three stages of motor learning, proposed by Fitts and Posner (84). The proficiency criteria include, in addition to time necessary to complete the task, also two motion efficacy metrics. Applying these metrics required an electromagnetic tracking device and thus increases the investment costs and reduces feasibility. Time measurements can be used singularly for continuous self-assessment during training, but for high-stakes assessment (certification), additional metrics or direct observation should be included to include information on errors or for eliminating deliberate shortcuts by the trainee (152).

Simulators serve as an adjunct to traditional training and should be regarded as an integrated part of a training programme for surgical residents. However, personal observations from several hospitals in Norway show that simulators, both VR-

simulators and box-trainers, are left unused if there is no systematic approach or administrative initiative encouraging simulation training. Only a few teaching hospitals have acquired VR-simulators, and there are no national guidelines or requirements for laparoscopic training with simulators of any kind. However, several teaching hospitals in Norway have already purchased a D-Box and there was a demand for establishing a curriculum for a systematic approach to laparoscopic training for surgical residents.

Motivation is important, both internal and external, to accomplish a successful training programme (71). Assessment of skills can act as a motivational factor, increasing the internal motivation for surgical training. If an assessment is aimed at changing behaviour (learning), it must be followed by an appropriate and objective feedback to the trainee. Studies show that feedback is essential and can lead to shortening of the learning curve and proficiency will thus be achieved faster and more efficiently (100,181).

The surgical community in Norway is reasonably small and transparent, compared to countries like USA and UK. All certification of surgeons are regulated and ratified by the government, and a national approach to certification of technical skills is feasible. Courses could be organised in regional centres, providing VR-simulators and box-trainer experiences, and then each surgical department should offer further and repeated training locally, using a nationally approved curricula. It is important to emphasise the need for providing allocated time, space, support staff and sufficient funds. A national approach could lead to an improved surgical training process, documentation of actual

skills, and consequently contribute to increased patient safety by reducing potential errors and patient suffering (182).

Simulation, in its present form, does not exclude the necessity for supervised training on patients, but represents a valuable adjunct. Simulation technology can however provide the tools to produce «pre-trained novices» that are already proficient in basic laparoscopic skills, before ever entering the OR.

6.0 Future perspectives

After these studies were finished, the continued development of the D-Box has resulted in a new design³, using a strong but lightweight, plastic material. The web camera has been replaced by a video camera, resulting in an improved screen resolution, and also omitting the need for a laptop or standalone computer. This reduces overall purchase cost and increases mobility. All task boxes are redesigned to match the new D-Box, and a few new task boxes have been introduced. Therefore, the results and criterion benchmarks established in the present studies does not necessarily apply to the newly designed D-Box. The new version thus demands a new and full validation process, including establishing new benchmark criteria based on expert performance on the latest version of D-Box. Our proposed curriculum must then also be reconfirmed/revalidated.

Further validation of the D-Box should include demonstrating whether the skills learned on the simulator transfers to a clinical setting, to the OR, and what impact this will have on clinical outcome.

In addition to the technical skills of laparoscopy, non-technical skills (theoretical knowledge, communication skills, decision making and leadership) also need to be addressed and included in a total curriculum for surgeons in training. Simulation methods are already available for all competencies.

³ <http://www.d-box.no/>

The technical advancements in the field of surgical simulation is fast and exciting and the imminent implementations include 3D-imaging, increasingly realistic haptic feedback and the incorporation of patient-specific clinical imagery (183). Future cooperation between clinicians, software and hardware developers will lead to training opportunities and simulation technologies beyond our present imagination.

References

- (1) Darzi A, Mackay S. Recent advances in minimal access surgery. *BMJ* 2002 Jan 5;324(7328):31-34.
- (2) Gurusamy K, Junnarkar S, Farouk M, Davidson BR. Meta-analysis of randomized controlled trials on the safety and effectiveness of day-case laparoscopic cholecystectomy. *Br J Surg* 2008 Feb;95(2):161-168.
- (3) Angenete E, Jacobsson A, Gellerstedt M, Haglind E. Effect of laparoscopy on the risk of small-bowel obstruction: a population-based register study. *Arch Surg* 2012 Apr;147(4):359-365.
- (4) Tiwari MM, Reynoso JF, High R, Tsang AW, Oleynikov D. Safety, efficacy, and cost-effectiveness of common laparoscopic procedures. *Surg Endosc* 2011 Apr;25(4):1127-1135.
- (5) Cuschieri A, Dubois F, Mouiel J, Mouret P, Becker H, Buess G, et al. The European experience with laparoscopic cholecystectomy. *Am J Surg* 1991 Mar;161(3):385-387.
- (6) Lacy AM, Garcia-Valdecasas JC, Delgado S, Castells A, Taura P, Pique JM, et al. Laparoscopy-assisted colectomy versus open colectomy for treatment of non-metastatic colon cancer: a randomised trial. *Lancet* 2002 Jun 29;359(9325):2224-2229.
- (7) Tsai CC, Lee SY, Huang FC. Laparoscopic versus open appendectomy in the management of all stages of acute appendicitis in children: a retrospective study. *Pediatr Neonatol* 2012 Oct;53(5):289-294.
- (8) Moazzez A, Mason RJ, Katkhouda N. Thirty-day outcomes of laparoscopic versus open appendectomy in elderly using ACS/NSQIP database. *Surg Endosc* 2012 Oct 6.
- (9) Cuschieri A, Berci G, McSherry CK. Laparoscopic cholecystectomy. *Am J Surg* 1990 Mar;159(3):273.
- (10) A prospective analysis of 1518 laparoscopic cholecystectomies. The Southern Surgeons Club. *N Engl J Med* 1991 Apr 18;324(16):1073-1078.
- (11) Bernard HR, Hartman TW. Complications after laparoscopic cholecystectomy. *Am J Surg* 1993 Apr;165(4):533-535.
- (12) de Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008 Jun;17(3):216-223.

- (13) Bakke KA. Slik sikkerhet tolereres ikke av andre. 2013; Available at: <http://www.dagensmedisin.no/nyheter/--slik-sikkerhet-tolereres-ikke-av-andre/>. Accessed 01/20, 2013.
- (14) To err is human: building a safer health system. Washington, D.C.; USA: National Academy Press; 2000.
- (15) Studdert DM, Mello MM, Gawande AA, Gandhi TK, Kachalia A, Yoon C, et al. Claims, errors, and compensation payments in medical malpractice litigation. N Engl J Med 2006 May 11;354(19):2024-2033.
- (16) Vincent C, Moorthy K, Sarker SK, Chang A, Darzi AW. Systems approaches to surgical quality and safety: from concept to measurement. Ann Surg 2004 Apr;239(4):475-482.
- (17) Smith R. All changed, changed utterly. British medicine will be transformed by the Bristol case. BMJ 1998 Jun 27;316(7149):1917-1918.
- (18) Ritchie WP, Jr. Basic certification in surgery by the American Board of Surgery (ABS): what does it mean? Does it have value? Is it relevant? A personal opinion. Ann Surg 2004 Feb;239(2):133-139.
- (19) Park AE, Witzke D. The surgical competence conundrum. Surgical endoscopy; Surgical endoscopy 2002 apr;16(4):555-557.
- (20) Halsted WS. The training of the surgeon. Bull Johns Hopkins Hosp 1904;15:267-275.
- (21) Helse- og omsorgsdepartementet. Spesialistregler i generell kirurgi. 2010; Available at: <http://legeforeningen.no/Emner/Spesialiteter/Generell-kirurgi/>. Accessed 01/20, 2013.
- (22) Helse- og omsorgsdepartementet. Spesialistregler for gastroenterologisk kirurgi. 2011; Available at: <http://legeforeningen.no/Emner/Spesialiteter/Gastroenterologisk-kirurgi/spesialistregler-i-gastroenterologisk-kirurgi/>. Accessed 01/20, 2013.
- (23) Helse- og omsorgsdepartementet. Generell bestemmelser for spesialistutdanningen. 2009; Available at: <http://legeforeningen.no/Emner/Andre-emner/Spesialistutdanning/Spesialisering/generelle-bestemmelser-for-spesialistutdanning-av-leger/>. Accessed 01/20, 2013.
- (24) McArdle CS, Hole DJ. Influence of volume and specialization on survival following surgery for colorectal cancer. Br J Surg 2004 May;91(5):610-617.
- (25) Marusch F, Koch A, Schmidt U, Pross M, Gastinger I, Lippert H. Hospital caseload and the results achieved in patients with rectal cancer. Br J Surg 2001 Oct;88(10):1397-1402.

- (26) Holm T, Johansson H, Cedermark B, Ekelund G, Rutqvist LE. Influence of hospital- and surgeon-related factors on outcome after treatment of rectal cancer with or without preoperative radiotherapy. *Br J Surg* 1997 May;84(5):657-663.
- (27) Smith JA, King PM, Lane RH, Thompson MR. Evidence of the effect of 'specialization' on the management, surgical outcome and survival from colorectal cancer in Wessex. *Br J Surg* 2003 May;90(5):583-592.
- (28) Pickersgill T. The European working time directive for doctors in training. *BMJ* 2001 Dec 1;323(7324):1266.
- (29) Ross DG, Harris CA, Jones DJ. A comparison of operative experience for basic surgical trainees in 1992 and 2000. *Br J Surg* 2002;89(Suppl 1):60.
- (30) Hutter MM, Kellogg KC, Ferguson CM, Abbott WM, Warshaw AL. The impact of the 80-hour resident workweek on surgical residents and attending surgeons. *Ann Surg* 2006 Jun;243(6):864-71; discussion 871-5.
- (31) Park A, Witzke D, Donnelly M. Ongoing deficits in resident training for minimally invasive surgery. *J Gastrointest Surg* 2002 May-Jun;6(3):501-7; discussion 507-9.
- (32) Carlin AM, Gasevic E, Shepard AD. Effect of the 80-hour work week on resident operative experience in general surgery. *Am J Surg* 2007 Mar;193(3):326-9; discussion 329-30.
- (33) Winslow ER, Bowman MC, Klingensmith ME. Surgeon workhours in the era of limited resident workhours. *Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract* 2004 Jan;198(1):111-117.
- (34) Chikwe J, de Souza AC, Pepper JR. No time to train the surgeons. *BMJ* 2004 Feb 21;328(7437):418-419.
- (35) Perissat J. Laparoscopic surgery: A pioneer's point of view. *World J Surg* 1999 Aug;23(8):863-868.
- (36) Litynski GS. Endoscopic surgery: the history, the pioneers. *World J Surg* 1999 Aug;23(8):745-753.
- (37) Semm K. Endoscopic appendectomy. *Endoscopy* 1983 Mar;15(2):59-64.
- (38) Mouret P. How I developed laparoscopic cholecystectomy. *Ann Acad Med Singapore* 1996 Sep;25(5):744-747.
- (39) Wright M, Turner D, Horbury C, Britain G, Berman G. Competence assessment for the hazardous industries. *Health and Safety Executive* 2003;086.
- (40) Flin R. Safe in Their Hands? Licensing and Competence Assurance for Safety-Critical Roles in High Risk Industries. 2005 10/2005.

(41) Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AH, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009 Jan 29;360(5):491-499.

(42) Wanzel KR, Ward M, Reznick RK. Teaching the surgical craft: From selection to certification. *Curr Probl Surg* 2002 Jun;39(6):573-659.

(43) Bell JA. Royal Air Force Selection Procedures. *Ann R Coll Surg Engl* 1988;70:270-275.

(44) Satava RM. Identification and reduction of surgical error using simulation. *Minim Invasive Ther Allied Technol* 2005;14(4):257-261.

(45) Bann S, Darzi A. Selection of individuals for training in surgery. *Am J Surg* 2005 Jul;190(1):98-102.

(46) Vapenstad C, Buzink SN. Procedural virtual reality simulation in minimally invasive surgery. *Surg Endosc* 2012 Sep 7.

(47) Halvorsen FH, Elle OJ, Fosse E. Simulators in surgery. *Minim Invasive Ther Allied Technol* 2005;14(4):214-223.

(48) Munz Y, Kumar BD, Moorthy K, Bann S, Darzi A. Laparoscopic virtual reality and box trainers: is one superior to the other? *Surgical Endoscopy and Other Interventional Techniques* 2004 Mar;18(3):485-494.

(49) Seymour NE. VR to OR: a review of the evidence that virtual reality simulation improves operating room performance. *World J Surg* 2008 Feb;32(2):182-188.

(50) Gurusamy K, Aggarwal R, Palanivelu L, Davidson BR. Systematic review of randomized controlled trials on the effectiveness of virtual reality training for laparoscopic surgery. *Br J Surg* 2008 Sep;95(9):1088-1097.

(51) Larsen CR, Soerensen JL, Grantcharov TP, Dalsgaard T, Schouenborg L, Ottosen C, et al. Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. *BMJ* 2009 May 14;338:b1802.

(52) Oostema JA, Abdel MP, Gould JC. Time-efficient laparoscopic skills assessment using an augmented-reality simulator. *Surg Endosc* 2008 Dec;22(12):2621-2624.

(53) Botden SM, Buzink SN, Schijven MP, Jakimowicz JJ. ProMIS augmented reality training of laparoscopic procedures face validity. *Simul Healthc* 2008 Summer;3(2):97-102.

(54) Pellen MG, Horgan LF, Barton JR, Attwood SE. Construct validity of the ProMIS laparoscopic simulator. *Surg Endosc* 2009 Jan;23(1):130-139.

- (55) Madan AK, Frantzides CT, Shervin N, Tebbit CL. Assessment of individual hand performance in box trainers compared to virtual reality trainers. *Am Surg* 2003 Dec;69(12):1112-1114.
- (56) Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg* 2002 Oct;236(4):458-63; discussion 463-4.
- (57) Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg* 2004 Feb;91(2):146-150.
- (58) Madan AK, Frantzides CT. Substituting virtual reality trainers for inanimate box trainers does not decrease laparoscopic skills acquisition. *JSLS* 2007 Jan-Mar;11(1):87-89.
- (59) Madan AK, Frantzides CT, Tebbit C, Shervin N. Self-reported versus observed scores in laparoscopic skills training. *Surgical Endoscopy and Other Interventional Techniques* 2005 May;19(5):670-672.
- (60) Hamilton EC, Scott DJ, Fleming JB, Rege RV, Laycock R, Bergen PC, et al. Comparison of video trainer and virtual reality training systems on acquisition of laparoscopic skills. *Surgical Endoscopy and Other Interventional Techniques* 2002 Mar;16(3):406-411.
- (61) Korndorffer JR, Jr, Clayton JL, Tesfay ST, Brunner WC, Sierra R, Dunne JB, et al. Multicenter Construct Validity for Southwestern Laparoscopic Videotrainer Stations. *J Surg Res* 2005 May 21.
- (62) Torkington J, Smith SG, Rees BI, Darzi A. Skill transfer from virtual reality to a real laparoscopic task. *Surgical Endoscopy and Other Interventional Techniques* 2001 Oct;15(10):1076-1079.
- (63) Bridges M, Diamond DL. The financial impact of teaching surgical residents in the operating room. *Am J Surg* 1999 Jan;177(1):28-32.
- (64) Youngblood PL, Srivastava S, Curet M, Heinrichs WL, Dev P, Wren SM. Comparison of training on two laparoscopic simulators and assessment of skills transfer to surgical performance. *J Am Coll Surg* 2005 Apr;200(4):546-551.
- (65) Diesen DL, Erhunmwunsee L, Bennett KM, Ben-David K, Yurcisin B, Ceppa EP, et al. Effectiveness of laparoscopic computer simulator versus usage of box trainer for endoscopic surgery training of novices. *J Surg Educ* 2011 Jul-Aug;68(4):282-289.
- (66) Brinkman WM, Havermans SY, Buzink SN, Botden SM, Jakimowicz JJ, Schoot BC. Single versus multimodality training basic laparoscopic skills. *Surg Endosc* 2012 Aug;26(8):2172-2178.

- (67) Sturm LP, Windsor JA, Cosman PH, Cregan P, Hewett PJ, Maddern GJ. A systematic review of skills transfer after surgical simulation training. *Ann Surg* 2008 Aug;248(2):166-179.
- (68) Glove PB. Webster's third new international dictionary of the English language, unabridged: with seven language dictionary. Chichago: Encyclopædia Britannica; 1981.
- (69) Ritchie WP,Jr. The measurement of competence. Current plans and future initiatives of the American Board of Surgery. *Bull Am Coll Surg* 2001 Apr;86(4):10-15.
- (70) Cuschieri A, Francis NK, Crosby J, Hanna GB. What do master surgeons think of surgical competence and revalidation? *Am J Surg* 2001 Aug;182(2):110-116.
- (71) Dankelman J, Chmarra MK, Verdaasdonk EG, Stassen LP, Grimbergen CA. Fundamental aspects of learning minimally invasive surgical skills. *Minim Invasive Ther Allied Technol* 2005;14(4):247-256.
- (72) Hasson HM. Core competency in laparoendoscopic surgery. *JSLs* 2006 Jan-Mar;10(1):16-20.
- (73) Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. *Br J Surg* 2004 Dec;91(12):1549-1558.
- (74) Gallagher AG, Richie K, McClure N, McGuigan J. Objective psychomotor skills assessment of experienced, junior, and novice laparoscopists with virtual reality. *World Journal of Surgery* 2001 Nov;25(11):1478-1483.
- (75) Gallagher AG, Ritter EM, Lederman AB, McClusky DA,3rd, Smith CD. Video-assisted surgery represents more than a loss of three-dimensional vision. *Am J Surg* 2005 Jan;189(1):76-80.
- (76) Ali MR, Mowery Y, Kaplan B, DeMaria EJ. Training the novice in laparoscopy. More challenge is better. *Surgical Endoscopy and Other Interventional Techniques* 2002 Dec;16(12):1732-1736.
- (77) Pearson AM, Gallagher AG, Rosser JC, Satava RM. Evaluation of structured and quantitative training methods for teaching intracorporeal knot tying. *Surgical Endoscopy and Other Interventional Techniques* 2002 Jan;16(1):130-137.
- (78) Hassan I, Gerdes B, Koller M, Dick B, Hellwig D, Rothmund M, et al. Spatial perception predicts laparoscopic skills on virtual reality laparoscopy simulator. *Childs Nerv Syst* 2007 Jun;23(6):685-689.
- (79) Peters JH, Fried GM, Swanstrom LL, Soper NJ, Sillin LF, Schirmer B, et al. Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. *Surgery* 2004 Jan;135(1):21-27.

- (80) Feldman LS, Sherman V, Fried GM. Using simulators to assess laparoscopic competence: ready for widespread use? *Surgery* 2004 Jan;135(1):28-42.
- (81) VAK learning styles. Available at:
<http://www.businessballs.com/vaklearningstylestest.htm>. Accessed February 14, 2012.
- (82) Kolb DA. *Experiential Learning: Experience as the source of learning and development*. 1st ed. New Jersey, USA: Prentice Hall Inc.; 1984.
- (83) Wentink M, Stassen LP, Alwayn I, Hosman RJ, Stassen HG. Rasmussen's model of human behavior in laparoscopy training. *Surg Endosc* 2003 Aug;17(8):1241-1246.
- (84) Fitts PM, Posner M.I. *Human Performance*. Belmont, CA, USA: Brooks & Cole; 1967.
- (85) Sherman V, Feldman LS, Stanbridge D, Kazmi R, Fried GM. Assessing the learning curve for the acquisition of laparoscopic skills on a virtual reality simulator. *Surgical Endoscopy and Other Interventional Techniques* 2005 May;19(5):678-682.
- (86) Taffinder NJ, Sutton C, Fishwick RJ, McManus IC, Darzi A. Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: results from randomised controlled studies using the MIST VR laparoscopic simulator. *Stud Health Technol Inform* 1998;50:124-130.
- (87) Deziel DJ, Millikan KW, Economou SG, Doolas A, Ko ST, Airan MC. Complications of laparoscopic cholecystectomy: a national survey of 4,292 hospitals and an analysis of 77,604 cases. *Am J Surg* 1993 Jan;165(1):9-14.
- (88) Aggarwal R, Ward J, Balasundaram I, Sains P, Athanasiou T, Darzi A. Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. *Ann Surg* 2007 Nov;246(5):771-779.
- (89) Dagash H, Chowdhury M, Pierro A. When can I be proficient in laparoscopic surgery? A systematic review of the evidence. *J Pediatr Surg* 2003 May;38(5):720-724.
- (90) Schijven MP, Jakimowicz J. The learning curve on the Xitact LS 500 laparoscopy simulator: profiles of performance. *Surgical Endoscopy and Other Interventional Techniques* 2004 Jan;18(1):121-127.
- (91) Van Sickle KR, Ritter EM, McClusky DA, 3rd, Lederman A, Baghai M, Gallagher AG, et al. Attempted establishment of proficiency levels for laparoscopic performance on a national scale using simulation: the results from the 2004 SAGES Minimally Invasive Surgical Trainer-Virtual Reality (MIST-VR) learning center study. *Surg Endosc* 2007 Jan;21(1):5-10.
- (92) Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. *Am J Surg* 2003 Feb;185(2):146-149.

- (93) Ahlberg G, Kruuna O, Leijonmarck CE, Ovaska J, Rosseland A, Sandbu R, et al. Is the learning curve for laparoscopic fundoplication determined by the teacher or the pupil? *Am J Surg* 2005 Feb;189(2):184-189.
- (94) Shapiro K, Patel S, Abdo Z, Ferzli G. Laparoscopic adjustable gastric banding: is there a learning curve? *Surgical Endoscopy and Other Interventional Techniques* 2004 Jan;18(1):48-50.
- (95) Tekkis PP, Senagore AJ, Delaney CP, Fazio VW. Evaluation of the learning curve in laparoscopic colorectal surgery: comparison of right-sided and left-sided resections. *Ann Surg* 2005 Jul;242(1):83-91.
- (96) Spencer F. Teaching and measuring surgical techniques: the technical evaluation of competence. *Bull Am Coll Surg* 1978;63(3):9-12.
- (97) Aggarwal R, Grantcharov T, Moorthy K, Milland T, Papasavas P, Dosis A, et al. An evaluation of the feasibility, validity, and reliability of laparoscopic skills assessment in the operating room. *Ann Surg* 2007 Jun;245(6):992-999.
- (98) Moorthy K, Munz Y, Dosis A, Bello F, Chang A, Darzi A. Bimodal assessment of laparoscopic suturing skills: Construct and concurrent validity. *Surgical Endoscopy and Other Interventional Techniques* 2004 Oct 13.
- (99) Gallagher AG, Satava RM, Shorten GD. Measuring surgical skill: a rapidly evolving scientific methodology. *Surg Endosc* 2013 May;27(5):1451-1455.
- (100) Halvorsen FH, Fosse E, Mjaland O. Unsupervised virtual reality training may not increase laparoscopic suturing skills. *Surg Laparosc Endosc Percutan Tech* 2011 Dec;21(6):458-461.
- (101) Fried GM, Feldman LS, Vassiliou MC, Fraser SA, Stanbridge D, Ghitulescu G, et al. Proving the value of simulation in laparoscopic surgery. *Ann Surg* 2004 Sep;240(3):518-25; discussion 525-8.
- (102) Reznick RK. Teaching and testing technical skills. *Am J Surg* 1993 Mar;165(3):358-361.
- (103) Martin J, Regehr G, Reznick RK, MacRae H, Murnaghan J, Hutchison C, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997 Feb;84(2):273-278.
- (104) Reznick RK, Regehr G, MacRae H, Martin J, McCulloch W. Testing technical skills outside the operating room: an innovative bench station examination. *Am J Surg* 1997 Mar;173(3):226-230.
- (105) Eubanks TR, Clements RH, Pohl D, Williams N, Schaad DC, Horgan S, et al. An objective scoring system for laparoscopic cholecystectomy. *J Am Coll Surg* 1999 Dec;189(6):566-574.

- (106) Dath D, Regehr G, Birch D, Schlachta C, Poulin E, Mamazza J, et al. Toward reliable operative assessment: the reliability and feasibility of videotaped assessment of laparoscopic technical skills. *Surgical Endoscopy and Other Interventional Techniques* 2004 Dec;18(12):1800-1804.
- (107) Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med* 1998 Sep;73(9):993-997.
- (108) Rosenbaum DA. Reaching and grasping. *Human Motor Control*. I ed. San Diego, USA: Academic Press Inc.; 1992. p. 197-225.
- (109) Mason JD, Ansell J, Warren N, Torkington J. Is motion analysis a valid tool for assessing laparoscopic skill? *Surg Endosc* 2013 May;27(5):1468-1477.
- (110) Hanna GB, Frank TG, Cuschieri A. Objective assessment of endoscopic knot quality. *Am J Surg* 1997 Oct;174(4):410-413.
- (111) Datta V, Mandalia M, Mackay S, Chang A, Cheshire N, Darzi A. Relationship between skill and outcome in the laboratory-based model. *Surgery* 2002 Mar;131(3):318-323.
- (112) Bann SD, Khan MS, Darzi AW. Measurement of surgical dexterity using motion analysis of simple bench tasks. *World J Surg* 2003 Apr;27(4):390-394.
- (113) Szalay D, MacRae H, Regehr G, Reznick RK. Using operative outcome to assess technical skill. *Am J Surg* 2000 Sep;180(3):234-237.
- (114) Satava RM, Cuschieri A, Hamdorf J, Metrics for Objective Assessment of Surgical Skills Workshop. Metrics for objective Assessment. *Surgical Endoscopy and Other Interventional Techniques* 2003 Feb;17(2):220-226.
- (115) Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surgical Endoscopy and Other Interventional Techniques* 2003 Oct;17(10):1525-1529.
- (116) Verdaasdonk EG, Stassen LP, Monteny LJ, Dankelman J. Validation of a new basic virtual reality simulator for training of basic endoscopic skills: the SIMENDO. *Surg Endosc* 2006 Mar;20(3):511-518.
- (117) Ayodeji ID, Schijven M, Jakimowicz J, Greve JW. Face validation of the Symbionix LAP Mentor virtual reality training module and its applicability in the surgical curriculum. *Surg Endosc* 2007 Sep;21(9):1641-1649.
- (118) Duffy AJ, Hogle NJ, McCarthy H, Lew JI, Egan A, Christos P, et al. Construct validity for the LAPSIM laparoscopic surgical simulator. *Surg Endosc* 2005 Mar;19(3):401-405.

- (119) Stefanidis D. Optimal acquisition and assessment of proficiency on simulators in surgery. *Surg Clin North Am* 2010 Jun;90(3):475-489.
- (120) Society of American Gastrointestinal and Endoscopic Surgeons. Fundamentals of Laparoscopic Surgery (FLS). 2013; Available at: <http://www.flsprogram.org>, 2013.
- (121) American Board of Surgeons. Training Requirements for General Surgery Certification. 2013; Available at: http://www.absurgery.org/default.jsp?certgsqe_training, 2013.
- (122) Xeroulis G, Dubrowski A, Leslie K. Simulation in laparoscopic surgery: a concurrent validity study for FLS. *Surg Endosc* 2009 Jan;23(1):161-165.
- (123) Sroka G, Feldman LS, Vassiliou MC, Kaneva PA, Fayez R, Fried GM. Fundamentals of laparoscopic surgery simulator training to proficiency improves laparoscopic performance in the operating room-a randomized controlled trial. *Am J Surg* 2010 Jan;199(1):115-120.
- (124) American College of Surgeons. Surgical Workforce in US. 2010; Available at: http://www.acshpri.org/documents/ACSHPRI_Surgical_Workforce_in_US_apr2010.pdf, 2013.
- (125) EAES. Laparoscopic Surgical Skills. 2013; Available at: <http://www.lss-surgical.eu>. Accessed 06/25, 2013.
- (126) Borel-Rinkes IH, Gouma DJ, Hamming JF. Surgical training in the Netherlands. *World J Surg* 2008 Oct;32(10):2172-2177.
- (127) van Dongen KW, Ahlberg G, Bonavina L, Carter FJ, Grantcharov TP, Hylander A, et al. European consensus on a competency-based virtual reality training program for basic endoscopic surgical psychomotor skills. *Surg Endosc* 2011 Jan;25(1):166-171.
- (128) Sutton C, McCloy R, Middlebrook A, Chater P, Wilson M, Stone R. MIST VR. A laparoscopic surgery procedures trainer and evaluator. *Stud Health Technol Inform* 1997;39:598-607.
- (129) Gallagher AG, Satava RM. Virtual reality as a metric for the assessment of laparoscopic psychomotor skills. Learning curves and reliability measures. *Surgical Endoscopy and Other Interventional Techniques* 2002 Dec;16(12):1746-1752.
- (130) Grantcharov TP, Rosenberg J, Pahle E, Funch-Jensen P. Virtual reality computer simulation. *Surgical Endoscopy and Other Interventional Techniques* 2001 Mar;15(3):242-244.
- (131) Gallagher AG, Lederman AB, McGlade K, Satava RM, Smith CD. Discriminative validity of the Minimally Invasive Surgical Trainer in Virtual Reality (MIST-VR) using criteria levels based on expert performance. *Surgical Endoscopy and Other Interventional Techniques* 2004 Apr;18(4):660-665.

- (132) Maithel S, Sierra R, Korndorffer J, Neumann P, Dawson S, Callery M, et al. Construct and face validity of MIST-VR, Endotower, and CELTS: are we ready for skills assessment using simulators? *Surg Endosc* 2006 Jan;20(1):104-112.
- (133) Chaudhry A, Sutton C, Wood J, Stone R, McCloy R. Learning rate for laparoscopic surgical skills on MIST VR, a virtual reality simulator: quality of human-computer interface. *Ann R Coll Surg Engl* 1999 Jul;81(4):281-286.
- (134) Brunner WC, Korndorffer JR, Jr, Sierra R, Dunne JB, Yau CL, Corsetti RL, et al. Determining standards for laparoscopic proficiency using virtual reality. *Am J Surg* 2005 Jan;71(1):29-35.
- (135) Torkington J, Smith SG, Rees BI, Darzi A. The role of the basic surgical skills course in the acquisition and retention of laparoscopic skill. *Surgical Endoscopy and Other Interventional Techniques* 2001 Oct;15(10):1071-1075.
- (136) Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. *BMJ* 2003 Nov 1;327(7422):1032-1037.
- (137) Moorthy K, Munz Y, Dosis A, Bello F, Darzi A. Motion analysis in the training and assessment of minimally invasive surgery. *Minim Invasive Ther Allied Technol* 2003 Jul;12(3):137-142.
- (138) Smith SG, Torkington J, Brown TJ, Taffinder NJ, Darzi A. Motion analysis. *Surgical Endoscopy and Other Interventional Techniques* 2002 Apr;16(4):640-645.
- (139) Dosis A, Aggarwal R, Bello F, Moorthy K, Munz Y, Gillies D, et al. Synchronized video and motion analysis for the assessment of procedures in the operating theater. *Arch Surg* 2005 Mar;140(3):293-299.
- (140) Leong JJ, Leff DR, Das A, Aggarwal R, Reilly P, Atkinson HD, et al. Validation of orthopaedic bench models for trauma surgery. *J Bone Joint Surg Br* 2008 Jul;90(7):958-965.
- (141) Saleh GM, Voyazis Y, Hance J, Ratnasothy J, Darzi A. Evaluating surgical dexterity during corneal suturing. *Arch Ophthalmol* 2006 Sep;124(9):1263-1266.
- (142) Hance J, Aggarwal R, Stanbridge R, Blauth C, Munz Y, Darzi A, et al. Objective assessment of technical skills in cardiac surgery. *Eur J Cardiothorac Surg* 2005 Jul;28(1):157-162.
- (143) Datta V, Mackay S, Mandalia M, Darzi A. The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. *Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract* 2001 Nov;193(5):479-485.
- (144) Datta V, Chang A, Mackay S, Darzi A. The relationship between motion analysis and surgical technical assessments. *Am J Surg* 2002 Jul;184(1):70-73.

- (145) Chmarra MK, Klein S, de Winter JC, Jansen FW, Dankelman J. Objective classification of residents based on their psychomotor laparoscopic skills. *Surg Endosc* 2010 May;24(5):1031-1039.
- (146) Grantcharov TP, Funch-Jensen P. Can everyone achieve proficiency with the laparoscopic technique? Learning curve patterns in technical skills acquisition. *Am J Surg* 2009 Apr;197(4):447-449.
- (147) Aggarwal R, Grantcharov TP, Eriksen JR, Blirup D, Kristiansen VB, Funch-Jensen P, et al. An evidence-based virtual reality training program for novice laparoscopic surgeons. *Ann Surg* 2006 Aug;244(2):310-314.
- (148) Aggarwal R, Crochet P, Dias A, Misra A, Ziprin P, Darzi A. Development of a virtual reality training curriculum for laparoscopic cholecystectomy. *Br J Surg* 2009 Sep;96(9):1086-1093.
- (149) Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. *Am J Surg* 2006 Jan;191(1):128-133.
- (150) Mackay S, Morgan P, Datta V, Chang A, Darzi A. Practice distribution in procedural skills training: a randomized controlled trial. *Surgical Endoscopy and Other Interventional Techniques* 2002 Jun;16(6):957-961.
- (151) Verdaasdonk EG, Stassen LP, van Wijk RP, Dankelman J. The influence of different training schedules on the learning of psychomotor skills for endoscopic surgery. *Surg Endosc* 2007 Feb;21(2):214-219.
- (152) Oropesa I, Sanchez-Gonzalez P, Lamata P, Chmarra MK, Pagador JB, Sanchez-Margallo JA, et al. Methods and tools for objective assessment of psychomotor skills in laparoscopic surgery. *J Surg Res* 2011 Nov;171(1):e81-95.
- (153) South-Eastern Norway Regional Health Authority, organization. 2013; Available at: <http://www.helse-sorost.no/omoss/english/Sider/organization.aspx>. Accessed 07/18, 2013.
- (154) Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. *N Engl J Med* 2003 Nov 27;349(22):2117-2127.
- (155) Hannan EL, Radzyner M, Rubin D, Dougherty J, Brennan MF. The influence of hospital and surgeon volume on in-hospital mortality for colectomy, gastrectomy, and lung lobectomy in patients with cancer. *Surgery* 2002 Jan;131(1):6-15.
- (156) Lavernia CJ, Guzman JF. Relationship of surgical volume to short-term mortality, morbidity, and hospital charges in arthroplasty. *J Arthroplasty* 1995 Apr;10(2):133-140.

- (157) Dardik A, Burleyson GP, Bowman H, Gordon TA, Williams GM, Webb TH, et al. Surgical repair of ruptured abdominal aortic aneurysms in the state of Maryland: factors influencing outcome among 527 recent cases. *J Vasc Surg* 1998 Sep;28(3):413-20; discussion 420-1.
- (158) Harling H, Bulow S, Moller LN, Jorgensen T, Danish Colorectal Cancer Group. Hospital volume and outcome of rectal cancer surgery in Denmark 1994-99. *Colorectal Dis* 2005 Jan;7(1):90-95.
- (159) Schrag D, Panageas KS, Riedel E, Cramer LD, Guillem JG, Bach PB, et al. Hospital and surgeon procedure volume as predictors of outcome following rectal cancer resection. *Ann Surg* 2002 Nov;236(5):583-592.
- (160) Simunovic M, To T, Baxter N, Balslem A, Ross E, Cohen Z, et al. Hospital procedure volume and teaching status do not influence treatment and outcome measures of rectal cancer surgery in a large general population. *J Gastrointest Surg* 2000 May-Jun;4(3):324-330.
- (161) Kapiteijn E, Marijnen CA, Colenbrander AC, Klein Kranenbarg E, Steup WH, van Krieken JH, et al. Local recurrence in patients with rectal cancer diagnosed between 1988 and 1992: a population-based study in the west Netherlands. *Eur J Surg Oncol* 1998 Dec;24(6):528-535.
- (162) Harmon JW, Tang DG, Gordon TA, Bowman HM, Choti MA, Kaufman HS, et al. Hospital volume can serve as a surrogate for surgeon volume for achieving excellent outcomes in colorectal resection. *Ann Surg* 1999 Sep;230(3):404-11; discussion 411-3.
- (163) Wibe A, Eriksen MT, Syse A, Tretli S, Myrvold HE, Soreide O, et al. Effect of hospital caseload on long-term outcome after standardization of rectal cancer surgery at a national level. *Br J Surg* 2005 Feb;92(2):217-224.
- (164) Martling A, Holm T, Rutqvist LE, Johansson H, Moran BJ, Heald RJ, et al. Impact of a surgical training programme on rectal cancer outcomes in Stockholm. *Br J Surg* 2005 Feb;92(2):225-229.
- (165) Heald RJ. Total mesorectal excision is optimal surgery for rectal cancer: a Scandinavian consensus. *Br J Surg* 1995 Oct;82(10):1297-1299.
- (166) Wibe A, Moller B, Norstein J, Carlsen E, Wiig JN, Heald RJ, et al. A national strategic change in treatment policy for rectal cancer--implementation of total mesorectal excision as routine treatment in Norway. A national audit. *Dis Colon Rectum* 2002 Jul;45(7):857-866.
- (167) Iacobucci G. Performance data on all surgeons in England will be published within two years. *BMJ* 2012 Dec 7;345:e8377.
- (168) Miskovic D, Ni M, Wyles SM, Kennedy RH, Francis NK, Parvaiz A, et al. Is competency assessment at the specialist level achievable? A study for the national

training programme in laparoscopic colorectal surgery in England. *Ann Surg* 2013 Mar;257(3):476-482.

(169) Stefanidis D, Scott DJ, Korndorffer JR, Jr. Do metrics matter? Time versus motion tracking for performance assessment of proficiency-based laparoscopic skills training. *Simul Healthc* 2009 Summer;4(2):104-108.

(170) Vapenstad C, Hofstad EF, Bo LE, Chmarra MK, Kuhry E, Johnsen G, et al. Limitations of haptic feedback devices on construct validity of the LapSim((R)) virtual reality simulator. *Surg Endosc* 2012 Dec 12.

(171) Gurusamy KS, Aggarwal R, Palanivelu L, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. *Cochrane Database Syst Rev* 2009 Jan 21;(1):CD006575. doi(1):CD006575.

(172) Nagendran M, Gurusamy KS, Aggarwal R, Loizidou M, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. *Cochrane Database Syst Rev* 2013 Aug 27;8:CD006575.

(173) Loukas C, Nikiteas N, Schizas D, Lahanas V, Georgiou E. A head-to-head comparison between virtual reality and physical reality simulation training for basic skills acquisition. *Surg Endosc* 2012 Sep;26(9):2550-2558.

(174) Lehmann KS, Ritz JP, Maass H, Cakmak HK, Kuehnafel UG, Germer CT, et al. A prospective randomized study to test the transfer of basic psychomotor skills from virtual reality to physical reality in a comparable training setting. *Ann Surg* 2005 Mar;241(3):442-449.

(175) Newmark J, Dandolu V, Milner R, Grewal H, Harbison S, Hernandez E. Correlating virtual reality and box trainer tasks in the assessment of laparoscopic surgical skills. *Am J Obstet Gynecol* 2007 Nov;197(5):546.e1-546.e4.

(176) Brinkman WM, Buzink SN, Alevizos L, de Hingh IH, Jakimowicz JJ. Criterion-based laparoscopic training reduces total training time. *Surg Endosc* 2012 Apr;26(4):1095-1101.

(177) Thompson JR, Leonard AC, Doarn CR, Roesch MJ, Broderick TJ. Limited value of haptics in virtual reality laparoscopic cholecystectomy training. *Surg Endosc* 2011 Apr;25(4):1107-1114.

(178) Chmarra MK, Dankelman J, van den Dobbelsteen JJ, Jansen FW. Force feedback and basic laparoscopic skills. *Surg Endosc* 2008 Oct;22(10):2140-2148.

(179) Strom P, Hedman L, Sarna L, Kjellin A, Wredmark T, Fellander-Tsai L. Early exposure to haptic feedback enhances performance in surgical simulator training: a prospective randomized crossover study in surgical residents. *Surg Endosc* 2006 Sep;20(9):1383-1388.

- (180) McGaghie WC, Draycott TJ, Dunn WF, Lopez CM, Stefanidis D. Evaluating the impact of simulation on translational patient outcomes. *Simul Healthc* 2011 Aug;6 Suppl:S42-7.
- (181) Strandbygaard J, Bjerrum F, Maagaard M, Winkel P, Larsen CR, Ringsted C, et al. Instructor feedback versus no instructor feedback on performance in a laparoscopic virtual reality simulator: a randomized trial. *Ann Surg* 2013 May;257(5):839-844.
- (182) Aggarwal R, Mytton OT, Derbrew M, Hananel D, Heydenburg M, Issenberg B, et al. Training and simulation for patient safety. *Qual Saf Health Care* 2010 Aug;19 Suppl 2:i34-43.
- (183) Willaert WI, Aggarwal R, Van Herzeele I, Cheshire NJ, Vermassen FE. Recent advancements in medical simulation: patient-specific virtual reality simulation. *World J Surg* 2012 Jul;36(7):1703-1712.

Papers I-IV