TOWARDS SAFER LAPAROSCOPIC SURGERY

TRAINING AND ASSESSMENT OF SURGICAL SKILLS AND EQUIPMENT HANDLING

DIEDERICK VAN HOVE

Towards safer laparoscopic surgery: training and assessment of surgical skills and equipment handling

Diederick van Hove



Title: Towards safer laparoscopic surgery: training and assessment of surgical skills and equipment handling.

Author: Diederick van Hove Lay-out & printing: Optima Grafische Communicatie, Rotterdam

Copyright 2013, P.D. van Hove, Delft, The Netherlands

All rights reserved. No part of this thesis may be reproduced by any means, or transmitted in any form without the written permission of the author or, when appropriate, of the publishers of the publications.

The research in this thesis was financially supported by a grant from the national healthcare company DSW (DSW, Schiedam, the Netherlands).

The printing of this thesis was financially supported by kind contributions from: Olympus Nederland B.V, ERBE Nederland B.V. and Simendo B.V.

ISBN: 978-94-6169-409-6

TOWARDS SAFER LAPAROSCOPIC SURGERY: TRAINING AND ASSESSMENT OF SURGICAL SKILLS AND EQUIPMENT HANDLING

Proefschrift

ter verkrijging van de graad van doctor aan de Technische Universiteit Delft, op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben, voorzitter van het College voor Promoties, in het openbaar te verdedigen op vrijdag 20 september 2013 om 12:30 uur door

Pieter Diederick VAN HOVE

Doctorandus in de Geneeskunde, arts geboren te 's-Gravenhage

DIT PROEFSCHRIFT IS GOEDGEKEURD DOOR DE PROMOTOREN:

Prof. dr. J. Dankelman Prof. dr. L.P.S. Stassen

SAMENSTELLING PROMOTIECOMMISSIE:

| voorzitter |
|---|
| Technische Universiteit Delft, promotor |
| Maastricht Universitair Medisch Centrum, promotor |
| Reinier de Graaf Gasthuis Delft |
| Technische Universiteit Delft |
| VU Medisch Centrum |
| Erasmus Medisch Centrum |
| Norwegian University of Science and Technology |
| |

TABLE OF CONTENTS

| CHAPTER 1 | General introduction | 9 |
|-----------|--|-----|
| 1.1 | Introduction | 10 |
| 1.2 | Problem statement and thesis objective | 14 |
| 1.3 | Outline of the thesis | 15 |
| CHAPTER 2 | Training and assessment of laparoscopic surgical skills | 19 |
| 2.1 | Objective assessment of technical surgical skills – review | 21 |
| 2.2 | An "intermediate curriculum" for advanced laparoscopic skills training with virtual reality simulation | 51 |
| CHAPTER 3 | Safety and competence: a comparison between healthcare and the petrochemical industry | 71 |
| 3.1 | Safety management, training and culture. What can we learn from the industry? | 73 |
| 3.2 | Training and assessment of equipment-related competence. Comparison of a petrochemical company and a hospital | 85 |
| CHAPTER 4 | Training for safe handling of technical laparoscopic equipment | 97 |
| 4.1 | Use of checklists for laparoscopic equipment in Dutch hospitals | 99 |
| 4.2 | Effect of basic laparoscopic skills courses on essential knowl- edge of equipment | 109 |
| 4.3 | Validation of an interactive simulation module to train the use of a laparoscopic insufflator | 123 |
| 4.4 | Development an validation of an interactive simulation module to train the use of an electrosurgical device | 135 |
| CHAPTER 5 | General conclusions, discussion and future perspectives | 145 |
| | English summary | 153 |
| | Nederlandse samenvatting | 159 |
| | List of publications | 164 |
| | List of congress presentations | 165 |
| | Curriculum vitae | 167 |
| | Dankwoord | 168 |

"Tell me and I will forget. Show me and I may remember. Involve me and I will understand." Confucius (450 BC)

CHAPTER 1

General introduction

1.1 INTRODUCTION

Patient safety and medical errors

Patient safety has received abundant attention in the past decade. With the publication of the American Institute of Medicine's report '*To Err is Human*' in 1999, it became obvious that healthcare is not infallible and that many errors are being made. The report revealed that 44.000 – 98.000 people died annually in the United States due to medical errors.¹ More studies worldwide have been performed and the median incidence of adverse events was calculated to be 9.2%. Of all adverse events, a median of 7.4% causes the death of a patient and a median percentage as high as 43.5% is preventable.² In The Netherlands a nationally conducted study found that 5.7% of all hospital-admitted patients in 2004 suffered from an adverse event, of which 8% was lethal and 40.4% was considered to be preventable.³ A subsequent study from the same institute 4 years later surprisingly showed a slight increase of these numbers.⁴ Although that might have been caused by increased reporting of events due to more attention for adverse events, these numbers suggest that patient safety has not majorly improved yet and that it is still an urgent topic.

The human factor in a system

Instinctively, one would argue that individuals in healthcare are to blame for the errors being made. However, the problem is not that simple, as was explained by the famous psychologist James Reason. He showed that individuals are part of a 'system' that can be 'diseased' and thereby can enable individuals to make errors. To explain this, Reason made 3 distinctions in errors. One of these distinctions is the one between active and latent failures and is most essential in understanding the system approach theory. Active failures are defined as unsafe acts at the sharp end of action (the operating room) and will have an immediate outcome. Conversely, latent failures have a delayed outcome which mostly is a result from decisions in higher layers of an organization.⁵ A system consists of multiple layers of which the bottom or inner layer represents the sharp end, where the actual action takes place. All other layers form the blunt end, representing the facilitating layers of an organization, like operating department management, technical services department and even higher layers like governmental policies (Fig.1). Mistakes at the blunt end of a system can eventually result in conditions at the sharp end which enable humans to make errors (a diseased system). The system approach theory assumes that humans make mistakes, but that a system should have defense mechanisms in it to prevent errors at the sharp end from taking place.⁶ Such defense mechanisms may consist of protocols, extra checkpoints (checklists) and (better) training of medical personnel.

General introduction



Figure 1 Onion model showing the different layers of a system. Adopted from: Dankelman⁶

Minimally invasive surgery

A substantial part of all adverse events, calculated to be 39.6%, is related to operations.² Therefore, surgery is one of the fields in which preventive measures can have a great effect on patient safety. Additionally, many adverse events related to surgery are ascribed to the increasing role of technology in the surgical field, especially accompanying the minimally invasive operative techniques.^{7, 8} Since its early use, already in 1985, this technique has been widely applied to many different operations and at present it has an undeniable position in surgical practice.⁹⁻¹¹

After the first years of application it was apparent that this technique requires fundamentally different skills compared to conventional surgery and therefore has a separate learning curve.^{12, 13} To start with, a 2D image of a 3D environment (the abdomen) is shown on a screen, which impedes depth perception.¹⁴⁻¹⁶ Secondly, long instruments are inserted through small incisions in the abdominal wall creating a pivot-point and these are in-between the surgeon's hands and the tissue that is operated on. This disturbs eyehand coordination, reduces the number of degrees-of-freedom and tactile feedback.¹⁷⁻¹⁹ But there are more skills required for performing minimally invasive surgery than just psychomotor skills. It has been observed that incidents happen with electronical operating equipment in 42 - 87% of laparoscopic operations.^{20, 21} The increasing amounts of technology required to practice new operating techniques, demand extra skills and knowledge from the operating room personnel which were previously not needed. In a report on minimally invasive surgery, the Dutch Healthcare Inspectorate stated that many unusual complications were reported related to this technology and that preventive measures were insufficient. According to this report, a required level of competence for practicing this technique was not defined. Moreover, training was stated to be unstructured and not uniform among different specialties and there was no system to assure competence by means of structured assessment.^{22, 23} Subsequent reports of the inspectorate confirmed the lack of training with technology among medical staff and additionally a lack of standardization in the whole operative process.^{24, 25} In reaction to these findings, the DHI demanded a discipline-exceeding approach to these problems. This included a demand for structured training programs for minimally invasive surgery and implementation thereof into specialist training programs and the use of video-analysis for objective assessment of surgical skills and certification. Furthermore, the use of methods for risk analysis and means for standardization of processes, like checklists, were strongly recommended.

Efficient training

As mentioned above, it is of great importance that skills are trained outside the operating room, before a surgeon performs surgery on real patients. In order to assure efficient training it is essential to distinct at first *what* exactly has to be trained and subsequently *how* that will be trained. If this is done consistently, it is likely to lead to the best result.

To make this distinction, several models can be used. One model of human behaviour has been described by Rasmussen.²⁶ Rasmussen divided human behaviour into 3 levels: skill-based behaviour, rule-based behaviour and knowledge-based behaviour.

- Skill-based behaviour: This level of behaviour represents highly automated task execution, which takes place without any conscious control. Examples are writing or playing a sport.
- Rule-based behaviour: This level of behaviour represents task execution according to
 a predefined order of steps which may have been derived from previous experience
 or other sources of information, for instance text books. It refers best to protocols or
 procedures which are to be followed in specified situations.
- Knowledge-based behaviour: This level of behaviour refers to situations for which no rules (procedures or protocols) exist. Therefore, first a goal has to be formulated by analysis of the situation and then a step-by-step plan has to be developed by careful consideration of multiple scenarios. This level of behaviour is associated with a higher level of abstraction as it requires a certain degree of 'mental modelling'. In surgery it is best exemplified by handling complications.

In his paper Rasmussen stated that, in order to make good man-machine interfaces, it is important to realize at which level of behaviour these are aimed. This remark can also be transferred to the development of training. However, the levels of behaviour

are dynamic. Knowledge-based behaviour, if used frequently, will become rule-based and this again can eventually even become skill-based. This transition to different levels refers closely to a model for learning skills, described by Fitts and Posner in 1967.²⁷ They described the *cognitive phase* in which a mental picture is formed of the different steps of a skill, followed by the *associative phase* in which these steps are practised repetitively and after which finally the *autonomous phase* is reached in which the skill has developed into an automated action. Remarkably, Fitts and Posner stated that not everyone will reach the autonomous state.

Transformation of surgical training

Transmission of knowledge and skills is essential in surgical training. Traditionally this used to take place according to the apprenticeship model (learning on the job), in which the surgical trainee initially performs small steps of an operation under continuous supervision of an experienced surgeon and is gradually allowed to expand his acts as the intensity of supervision decreases.²⁸

Under the influence of numerous aspects, surgical training programs have changed. As mentioned above, simply more severe demands have been put up by authorities. However, there are other reasons why surgical training was forced to evolve. The medical trainees' workweek has been reduced from 100 to 80 hours in the United States and to 48 hours in Europe, while the number of years to become a medical specialist has remained equal.^{29, 30} This asks for more efficient use of the time spent in the operating room. Additionally, influenced by the discussion on medical errors, ethical reference within doctor-patient relationships has changed and it is not considered acceptable anymore that a surgeon proceeds through his learning curve purely at the cost of patients serving as 'training objects'.

Therefore, the apprenticeship model has been accompanied by competency-based training programs. These training programs are largely based on the CANMEDS framework, developed by the Royal College of Physicians and Surgeons of Canada.³¹ The CANMEDS framework is based on the fact that being a good physician (surgeon) is not one single competency, but a mix of multiple competencies. The CANMEDS framework has defined seven competencies (medical expert, professional, communicator, collaborator, manager, health advocate, scolar) which all need separate training to become a good physician and/or surgeon. Moreover, the CANMEDS framework also consists of methods for assessment of all defined competencies. Another important aspect of the framework is that the trainee himself is responsible for receiving adequate training programs have been rewritten, based on the CANMEDS framework.

1.2 PROBLEM STATEMENT AND THESIS OBJECTIVE

The introduction of minimally invasive surgery, together with the attention for medical errors and the new structure of surgical training programs has resulted in the development of multiple methods for skills training and assessment of laparoscopic surgery.³² This confluence of developments in surgery has revealed new problems. In this thesis a number of these problems will be discussed.

To start with, newly developed methods for training and assessment of skills do not always get implemented into training programs. One of the reasons for this is that it is unclear what should be the exact role of the different methods within a program. Validity of many training methods has been tested, including virtual reality (VR) simulators. VR simulators offer advantages like the ability of endless use and objective scoring without the need of an observer. Training with VR simulators has been validated for basic laparoscopic skills, like eye-hand coordination, and was even proven to transfer to the operating room.³³ To train more advanced surgical skills other methods, like conventional box trainers or animal cadavers, are required. However, the development of VR simulators has improved to increasingly realistic environments and several manufacturers now offer simulators featuring a virtual abdomen and force feedback in order to expand the possibilities of VR simulation.³⁴ Nonetheless, it is doubtful whether such simulators are realistic enough to train all steps of an operation sufficiently. Likewise, as data of tactile feedback of real living human tissue is lacking, force feedback in these simulators may not be truly realistic.³⁵ Some studies even suggest that unrealistic force feedback can result in a negative learning effect.³⁶ Therefore, currently VR is valid for training basic laparoscopic skills but other modalities are used for more advanced skills training. Whether VR simulation can be used for this as well should be investigated.

Secondly, methods for assessment of surgical skills are increasingly desired by controlling instances to be used for certification of surgeons. The Dutch Healthcare Inspectorate has suggested letting laparoscopic surgeons judge each others' skills by blinded videos of laparoscopic operations. However, it is unknown if any assessment method is currently suitable enough to be used for this purpose and surgical societies are therefore reluctant to introduce an official method for skills assessment and certification.

Finally, the increasing role of technology in surgery asks for attention for equipmentrelated errors. Observational studies showed that incidents often happen with the equipment during laparoscopic surgery.^{20, 21} These incidents vary from equipment not being present, to faulty connections or settings. A checklist specifically focussing on this equipment can reduce the amount of incidents by 50%.³⁷ Although the use of checklists is gaining popularity, it is unknown to which extent checklists specifically aimed at equipment are currently being used in hospitals. Additionally, much is unknown about training with surgical equipment. The subject is more or less ignored in literature and often the organisation is left to the responsibility of departments themselves. Therefore it is not transparent how such training is arranged and what its effect is. Nonetheless, the numbers of observed incidents with the equipment suggest that equipment-related training should be taken seriously and could be more effective.

Thesis objective

The objective of this thesis is to improve laparoscopic surgical skills training and assessment, with an emphasis on equipment-related safety and competence.

1.3 OUTLINE OF THE THESIS

This thesis is subdivided into 3 parts.

Training and assessment of laparoscopic surgical skills

Chapter 2.1 provides an overview of all methods currently available for objective assessment of technical surgical skills and their evidence.

Chapter 2.2 describes the validation of an "intermediate curriculum" for training of advanced laparoscopic skills on a virtual reality simulator.

Safety and competence: a comparison between healthcare and the petrochemical industry

Chapter 3.1 describes the differences and similarities between healthcare and the petrochemical industry with regard to a safety management system and safety culture. Chapter 3.2 describes the comparison of a hospital and a petrochemical company with regard to equipment-related training and assessment.

Training for the safe handling of technical laparoscopic equipment

Chapter 4.1 provides an overview of the current use of checklists for laparoscopic operating equipment in Dutch hospitals and the training of their personnel with this equipment. Chapter 4.2 evaluates the effect of 3 basic laparoscopic skills courses on participants' knowledge of technical laparoscopic equipment.

Chapter 4.3 describes the development and construct and face validation of an interactive, web-based, simulation module for the use of a laparoscopic insufflator.

Chapter 4.4 describes the development and evaluation of the training effect of an interactive, web-based simulation module for the use of an electrosurgical device.

Chapter 5 provides a general discussion on this thesis and possible directions for future research. Finally, conclusions to this thesis are given.

REFERENCES

- 1. Kohn LT CJ, Donaldson MS, editor. To err is human: building a safer health system. Washington DC: National Academy Press; 1999.
- 2. 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;17:216-23.
- de Bruijne MC, Zegers M, Hoonhout LHF, Wagner C. Onbedoelde schade in Nederlandse ziekenhuizen. Dossieronderzoek van ziekenhuisopnames in 2004. Utrecht, The Netherlands. 2007. p. 1-125.
- Langelaan M, Baines RJ, Broekens MA, Siemerink KM, van de Steeg L, Asscheman H, et al. Monitor zorggerelateerde schade 2008: Dossieronderzoek in Nederlandse ziekenhuizen. Utrecht, The Netherlands. 2010. p. 1-120 http://www.nivel.nl/pdf/Rapport-zorggerelateerde-schade.pdf.
- 5. Reason J. Understanding adverse events: human factors. Qual Health Care. 1995;4:80-9.
- 6. Dankelman J, Grimbergen CA. Systems approach to reduce errors in surgery. Surg Endosc. 2005;19:1017-21.
- Cuschieri A. Whither minimal access surgery: tribulations and expectations. Am J Surg. 1995;169:9-19.
- 8. Panesar SS, Salvilla SA, Patel B, Donaldson SL. Laparoscopic cholecystectomy: device-related errors revealed through a national database. Expert Rev Med Devices. 2011;8:555-60.
- Buunen M, Veldkamp R, Hop WC, Kuhry E, Jeekel J, Haglind E, et al. Survival after laparoscopic surgery versus open surgery for colon cancer: long-term outcome of a randomised clinical trial. Lancet Oncol. 2009;10:44-52.
- 10. Draaisma WA, Rijnhart-de Jong HG, Broeders IA, Smout AJ, Furnee EJ, Gooszen HG. Five-year subjective and objective results of laparoscopic and conventional Nissen fundoplication: a randomized trial. Ann Surg. 2006;244:34-41.
- Liem MS, van der Graaf Y, van Steensel CJ, Boelhouwer RU, Clevers GJ, Meijer WS, et al. Comparison of conventional anterior surgery and laparoscopic surgery for inguinal-hernia repair. N Engl J Med. 1997;336:1541-7.
- 12. Moore MJ, Bennett CL. The learning curve for laparoscopic cholecystectomy. The Southern Surgeons Club. Am J Surg. 1995;170:55-9.
- Figert PL, Park AE, Witzke DB, Schwartz RW. Transfer of training in acquiring laparoscopic skills. J Am Coll Surg. 2001;193:533-7.
- Crosthwaite G, Chung T, Dunkley P, Shimi S, Cuschieri A. Comparison of direct vision and electronic two- and three-dimensional display systems on surgical task efficiency in endoscopic surgery. Br J Surg. 1995;82:849-51.
- 15. Cuschieri A. Visual Displays and Visual Perception in Minimal Access Surgery. Semin Laparosc Surg. 1995;2:209-14.
- 16. Crothers IR, Gallagher AG, McClure N, James DT, McGuigan J. Experienced laparoscopic surgeons are automated to the "fulcrum effect": an ergonomic demonstration. Endoscopy. 1999;31:365-9.
- 17. Bholat OS, Haluck RS, Murray WB, Gorman PJ, Krummel TM. Tactile feedback is present during minimally invasive surgery. J Am Coll Surg. 1999;189:349-55.
- 18. Gallagher AG, McClure N, McGuigan J, Ritchie K, Sheehy NP. An ergonomic analysis of the fulcrum effect in the acquisition of endoscopic skills. Endoscopy. 1998;30:617-20.
- 19. Tholey G, Desai JP, Castellanos AE. Force feedback plays a significant role in minimally invasive surgery: results and analysis. Ann Surg. 2005;241:102-9.

General introduction

- 20. Courdier S, Garbin O, Hummel M, Thoma V, Ball E, Favre R, et al. Equipment failure: causes and consequences in endoscopic gynecologic surgery. J Minim Invasive Gynecol. 2009;16:28-33.
- 21. Verdaasdonk EG, Stassen LP, van der Elst M, Karsten TM, Dankelman J. Problems with technical equipment during laparoscopic surgery. An observational study. Surg Endosc. 2007;21:275-9.
- 22. Inspectie voor de Gezondheidszorg (IGZ). Risico's minimaal invasieve chirurgie onderschat. Den Haag: Inspectie voor de Gezondheidszorg (IGZ). 2007. p. 1-93.
- 23. Stassen LP, Bemelman WA, Meijerink J. Risks of minimally invasive surgery underestimated: a report of the Dutch Health Care Inspectorate. Surg Endosc. 2009;24:495-8.
- 24. Inspectie voor de Gezondheidszorg (IGZ). Standaardisatie onmisbaar voor risicovermindering in operatief proces. Den Haag: Inspectie voor de Gezondheidszorg (IGZ). 2008. p. 1-80.
- 25. Inspectie voor de Gezondheidszorg (IGZ). Risico's van medische technologie onderschat. Den Haag: Inspectie voor de Gezondheidszorg (IGZ). 2008. p. 1-81.
- 26. Rasmussen J. Skills, Rules and Knowledge; SIgnals, Signs and Symbols and Other Distinctions in Human Performance Models. IEEE Trans Syst Man Cybern. 1983;13:257-66.
- 27. Fitts P.M, Posner M.I. Human Performance. Belmont, CA: Brooks and Cole; 1967.
- 28. Halsted WS. The training of the surgeon. Bulletin of the John Hopkins Hospital. 1904;15:267-75.
- 29. Accreditation Council for Graduate Medical Education (ACGME). Common program requirements July 1, 2011. 2011.p.1-19
- 30. European Parliament and Council. Directive 2000/34/EC of 22 June 2000. 2000.p.41-5
- 31. Frank JR. The CanMEDS 2005 Physician Competency Framework. Better standards. Better physicians. Better care. Ottawa, Canada. 2005.
- 32. Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. Br J Surg. 2004;91:1549-58.
- 33. Seymour NE. VR to OR: A review of the evidence that virtual reality simulation improves operating room performance. World JSurg. 2008;32:182-8.
- 34. Ikonen TS, Antikainen T, Silvennoinen M, Isojarvi J, Makinen E, Scheinin TM. Virtual reality simulator training of laparoscopic cholecystectomies - a systematic review. Scand J Surg. 2012;101:5-12.
- 35. Coles TR, Meglan D, John NW. The role of haptics in medical training simulators: A survey of the state-of-the-art. IEEE Transactions on Haptics. 2011;4:51-66.
- 36. Chmarra MK, Dankelman J, van den Dobbelsteen JJ, Jansen FW. Force feedback and basic laparoscopic skills. Surg Endosc. 2008;22:2140-8.
- 37. Verdaasdonk EG, Stassen LP, Hoffmann WF, van der Elst M, Dankelman J. Can a structured checklist prevent problems with laparoscopic equipment? Surg Endosc. 2008;22:2238-43.

TRAINING AND ASSESSMENT OF LAPAROSCOPIC SURGICAL SKILLS

CHAPTER 2.1

Objective assessment of technical surgical skills - review

P.D. van Hove^{1,2}, G.J.M. Tuijthof^{1,3}, E.G.G. Verdaasdonk², L.P.S. Stassen⁴, J. Dankelman¹

1: Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands 2: Department of Surgery, Reinier de Graaf Group, Delft, the Netherlands 3: Department of Orthopaedic Surgery, Amsterdam Medical Centre, Amsterdam, the Netherlands 4: Department of Surgery, Maastricht University Medical Center, Maastricht, the Netherlands

Published in: Br J Surg. 2010 Jul;97(7):972-87

ABSTRACT

Background At present, surgeons are increasingly scrutinized for their performance and objective assessment methods for technical skills have gained interest. The purpose of this study is to review all evidence for these methods, in order to provide a guideline for use in clinical practice.

Methods A systematic search was performed using Pub Med and Web of Science for studies addressing validity and reliability of methods for objective assessment within surgery and gynaecology only. The studies were assessed according to the Oxford Centre for Evidence-based Medicine Levels of Evidence.

Results In total 104 studies were included, of which 20 studies (19.2%) had a level of evidence 1b or 2b. In 28 studies (26.9%), the assessment method was used in the operating room. Virtual reality simulators and Objective Structured Assessment of Technical Skills (OSATS) have been studied most. Although OSATS is seen as the gold standard for skills assessment, only 7 studies with a low level of evidence address its use in the OR.

Conclusion Based on currently available evidence, most methods for skills assessment are considered valid for feedback or measuring progress of skills, but few can be used for examination or credentialing. The purpose of assessment determines the choice of a proper method.

INTRODUCTION

Traditionally, surgical skills have been assessed in the operating room by supervision and feedback^{1, 2}. However, this method has been criticized for being too subjective and not representing the actual level of skills³. At present, there is an increasing demand from society, followed by governments and insurance companies, for clear and transparent quality measurements of healthcare, and surgeons and surgical trainees are increasingly scrutinized for their performance⁴⁻⁶. Additionally, new techniques, such as minimal access surgery, require new skills, which have different learning curves and require different training methods outside the operating room^{1, 7, 8}. These developments have resulted in an increased interest for objective assessment methods for surgical skills, which are currently used in surgical residency programs for assessing performance of trainees and serving as feedback with training. Moreover, these methods are desired to be applied as tools for examination in, for instance, different stages of residency. Likewise, governments are planning to use such methods for assessing competence of practicing laparoscopic surgeons in order to use it for credentialing^{1, 3, 4, 9, 10}.

In the past years, different methods for objective assessment of surgical skills have been developed and studies addressing validity and reliability are abundant^{5, 9, 11-14}. However, so far, methods for objective assessment have not been widely adopted into clinical practice. This is most likely caused by a lack of expertise, the proper infrastructure for implementation and cost, but it could also be that educators are hesitant to use them because it is not yet fully defined how and where these methods can be used. Consequently, bad choices could be made, resulting in implementing inappropriate methods.

There is lack of a good overview of current methods for objective assessment and their capabilities. Published reviews only tend to sum up and describe different methods^{9, 11, 13}. Therefore, the purpose of this study is to provide a review of the current evidence for objective assessment methods for technical surgical skills.

METHODS AND MATERIALS

A systematic search of the literature was performed, using Pub Med and Web of Science, for studies concerning validity or reliability of methods for assessment of technical surgical skills. The following query was used: "(surgical OR operative OR laparoscopic OR technical) AND (skills OR competence) AND assessment". Studies were included that addressed assessment methods which are applicable in or outside the OR and concern open surgery or laparoscopy in the domain of general surgery and gynaecology. Studies concerning other domains were not included. Only English language studies were included. Studies addressing the validity of specific bench models or simulator tasks, so-called settings, were excluded. However, studies were included when they used nonvalidated methods for rating a bench task as such studies contribute to validating the method as well as the bench task. Reviews and congress abstracts were excluded.

All studies were divided into separate categories based on the type of assessment method. Some studies discussed more than one assessment method and therefore contribute to more than one category. The following categories were defined: procedure-specific checklists, global rating scales, motion analysis, virtual reality simulators, video-assessment and miscellaneous. Extra categories were defined for *Objective Structured Assessment of Technical Skills* (OSATS) and *Fundamentals of Laparoscopic Skills* (FLS) manual skills test, because these two methods have both been studied extensively and are used in clinical practice.

All studies were rated according to the Oxford Centre for Evidence-based Medicine Levels of Evidence¹⁵ using the category for 'diagnostic studies', as validating studies can best be compared with diagnostic studies. Results and evidence for each category are summarized in a separate table, and the most important findings are discussed in separate sections.

Validity, reliability and types of assessment

Validity is defined as 'the property of being true, correct and in conformity with reality'¹⁶ and is subdivided in different levels: face validity, content validity, construct validity, concurrent validity, and predictive validity. Face validity addresses users' opinion about the functionality and realism of a test. Content validity refers to whether the content of a test is suited to measure what it is supposed to measure. Construct validity refers to whether a test indeed measures the trait it is supposed to measure; in this case, technical surgical skill. Discriminate validity is a variant of construct validity and requires a test to discriminate even more specifically, for instance between different experts. Concurrent validity is an expression of the comparison of a test to a gold standard, or another test which measures the same trait. Predictive validity refers to the extent to which a test predicts future performance^{16, 17}.

Reliability refers to whether a test is consistent in its outcome. Evidently, this also affects the validity of a test. Frequently used items for reliability are internal consistency, inter-rater reliability and inter-test (test-retest) reliability. Internal consistency reflects the correlation between different items of a test and how these items contribute to the outcome of the test. Inter-rater reliability refers to the agreement of the scores of 2 or more raters testing the same subject. This is best tested with raters who are unaware of the subject's training level and identity (i.e., blinded raters). Inter-test reliability refers to the agreement of scores when the same test is taken twice¹⁷. Reliability is represented by a reliability coefficient, which ranges from 0 to 1.0. Generally, 0.8 is accepted as a threshold for good reliability¹⁸ and was therefore adhered to in this study as well.

Finally, assessment can be either formative or summative. Formative assessment aims at development by monitoring a trainee's development progress and giving structured feedback. When an assessment method is to be used for formative assessment, it should be able to identify different levels of performance (i.e., construct validity). A summative assessment is to be used, at its highest ability, for selection and therefore needs predefined consequences of its outcome. For instance, an exam can be passed or failed and there is a pre-set threshold which has to be reached for passing. Summative assessment would be required for credentialing. Higher standards for construct validity and reliability are required with this form of assessment than with formative assessment. Moreover, clear cut-off values have to be defined adherent to the predefined consequences and, ideally, sensitivity and specificity of these values should be tested.

RESULTS

The search resulted in 931 unique studies, of which first all titles were assessed for relevancy. After title assessment, 257 studies were selected. Of these studies, all abstracts were read and assessed for inclusion criteria by 2 authors. Discrepancies were solved by discussion. After abstract assessment 104 studies were considered for further analysis. Twenty-two studies were excluded after studying the full text and 22 new relevant references were identified from reference lists. In total, 104 studies were left for review (see Figure 1). Of these 104 studies, 20 studies (19.2%) offered level 1b or 2b evidence. Only in 28 studies (26.9%) the assessment method was used in the operating room.

Procedure-specific checklists

Procedure-specific checklists are specifically designed for different procedures and usually follow subsequent steps of a procedure which are scored. Nine studies were identified concerning 8 procedure-specific checklists¹⁹⁻²⁷. Levels of evidence extended from 2b to 4. Five checklists were used in the OR^{19-22, 24, 27}, of which 3 were designed for laparoscopic cholecystectomies (LC)^{19, 21, 22, 24}. Two of these five checklists were used in combination with video registration. (See Table 1 for details).

The only studies with a high level of evidence (2b) are by Sarker²¹ and Eubanks¹⁹. Three checklists to be used for assessment of LC, were designed by Sarker. Two of these checklists showed construct validity. Inter-rater reliability was above the cut-off value of 0.8, meaning their reliability is good. Another checklist for LC was designed by Eubanks¹⁹ and moderate correlation with experience and reasonable to good inter-rater reliability were found. The same checklist was studied by Aggarwal²⁴ and worse results were found. All other studies have lower levels of evidence, either due to non-consecutive cohorts (level 3b), which could imply a selection bias of participants, or unblinded raters (level 4).



Figure 1 Flowchart of structured selection of eligible articles

Global rating scales

Global rating scales (GRS) are used to rate more general skills, which are applicable to all surgical procedures and thus not procedure-specific. Eleven studies were identified concerning 8 different GRS^{7, 28-37}. Two GRS were studied in a lab-setting, the other 6 in the OR. Only 2 studies consisted of level 1b or 2b evidence^{28, 37}, while all other studies consisted of level 4 evidence. (See Table 2 for details).

Except for two scales ^{28, 34}, all were used to assess live operations. Every GRS has been studied with different operations, except for those of Sidhu³⁴ and the Global Operative Assessment of Laparoscopic Skills (GOALS)^{33, 35-37}. These were respectively tested for laparoscopic colectomy in a porcine model, and for laparoscopic cholecystectomy and appendectomy in humans.

A study by Bramson²⁸ is one of the two studies offering a high level of evidence. Development of a GRS for use with small tasks on animal tissue lab-models is described, good correlation with surgical skills (estimated by questionnaire) is established and reliability is above 0.8.

The only GRS which has been tested in multiple studies is the GOALS. Four studies addressed this GRS for laparoscopy ^{33, 35-37}. First, it was developed by Vassiliou³⁵, who ap-

| Ref | Setting | Mode | le Checklist Level | | N | Construct validity | Reliability | | |
|-----|---------|-------------------------|--|----|-----|-----------------------|-------------------------|----------------------------|--|
| | | | | | | | Internal consistency | Inter-rater reliability | |
| 19 | OR | Video | Checklist score and error | 2b | 30 | 0.50ª | NA | 0.74-0.96 | |
| 24 | OR | Video | score for LC | 3b | 47 | No | NA | 0.58 | |
| 21 | OR | Video&Live ^b | Technical and technological skills for LC | 2b | 100 | Yes / No ^c | NA | >0.8 | |
| 22 | OR | Video | Generic & specific skills | 3b | 50 | Yes | NA | >0.8 | |
| 24 | OR | Video | for LC | 3b | 47 | No | | 0.62 | |
| 23 | Lab | Video | Checklist for 10-stations | 3b | 21 | Yes ^d | NA | 0.78 | |
| 25 | Lab | Video | Checklist for intracorporeal suturing | 3b | 26 | Yes | NA | 0.90 | |
| 26 | Lab | Video | Rating for low-anterior resection and Nissen fundoplication on a pig | 3b | 29 | No | NA | 0.73 | |
| 27 | OR | Video | Checklist for tubal banding | 4 | 23 | No | NA | 0.007-0.88 | |
| 20 | OR | Live | Rating different key procedures | 4 | 300 | Yes | >0.90 | NA | |
| | | | | | | | | | |

 Table 1
 Specifications and most important results for all included studies addressing procedurespecific checklists.

NA Not addressed

LC Laparoscopic cholecystectomy

a Coefficient for correlation of checklist score with experience

b Technical skills were assessed from recorded video. Technological OR equipment skills were assessed live.

c Technical skills were divided in generic and specific. Construct validity was not established for specific technical skills checklist.

d Construct validity for 6 of 10 work stations

plied it to the dissection phase of an LC. It appeared to be highly reliable and construct validity was established for all separate domains. Second, Gumbs³³ investigated whether GOALS would also be applicable to a total LC and even other laparoscopic operations. Ninety-four ratings of residents performing an LC or laparoscopic appendectomy were assembled and construct validity was established. In another study by McCluney³⁶, predictive and concurrent validity were established by comparing the GOALS-score with the score on the FLS simulator (correlation 0.77). Eventually, in a study by Chang³⁷, results were studied for 10 blinded observers, rating videos of a novice and an expert while performing an LC. Construct validity was established for 4 of 5 domains and high inter-rater reliability was found with level of evidence 1b, but only two video tapes were rated. With this, reliability of GOALS and the fact that it might be useful for video assessment, were proved. Although all 4 studies show consistently good results for GOALS, 3 of them consist of level 4 evidence, because the raters were not blinded.

| Ref | lef Setting Mode | | Global rating scale | Level | N | Construct validity | Relia | bility |
|-----|------------------|-------|--|-------|-----|-----------------------|-------------------------|----------------------------|
| | | | | | | | Internal consistency | Inter-rater reliability |
| 28 | Lab | Live | 10-item rating scale for basic skills | 2b | 65 | r=0.86ª | 0.84 | 0.83 |
| 29 | OR | Live | General and case-specific skills scale | 4 | 362 | Yes | >0.80 | |
| 7 | OR | Live | Rating scale for 5 competencies | 4 | 40 | r=0.96 ^b | | >0.96 |
| 30 | OR | Live | Rating scale for 3 competencies | 4 | 4 | NA | | 0.82 |
| 35 | OR | Live | Global operative | 4 | 21 | Yes | 0.91-0.94 | 0.82-0.89 |
| 33 | OR | Live | assessment of | 4 | 94 | Yes | | |
| 36 | OR | Live | laparoscopic skills. | 4 | 40 | Yes | | |
| 37 | OR | Video | (GOALS) | 1b | 2 | Yes ^c | | 0.87-0.93 |
| 31 | OR | Live | Modification of OSATS | 4 | 7 | Yes | >0.91 | |
| | | | and GUALS | | | 1=0.943 | | |
| 34 | Lab | Video | Modification of OSATS and GOALS | 4 | 22 | NA | >0.88 | 0.76 |
| 32 | OR | Live | Telephonic rating scale | 4 | 993 | Yes | | 0.28 |
| | | | | | | | | |

 Table 2
 Specifications and most important results for all included studies addressing global rating scales.

NA Not addressed a Correlation with surgical

a Correlation with surgical skillsb Correlation with faculty ratings

c For 4 of 5 domains.

d Correlation with postgraduate year (PGY)

Objective Structured Assessment of Technical Skills (OSATS)

Objective Structured Assessment of Technical Skills (OSATS) is one of the first methods designed for objective skills assessment. It is also the instrument which has been studied most extensively and is one of the few instruments that is actually used in clinical practice. It consists of a GRS and a procedure-specific checklist. Originally, it was designed for use in lab settings, but it is now also used in the OR.

Twenty-six studies were identified that addressed OSATS^{26, 38-62}. Nineteen studies cover OSATS in the lab-setting ^{26, 38, 41-47, 49-54, 57-59, 61} and seven in the OR^{39, 40, 48, 55, 56, 60, 62}. (See Table 3 for details).

In total, construct validity was established in 18 studies, internal consistency was above 0.8 in 12 studies and inter-rater reliability was above 0.8 in 10 studies (see Table 3). For use in a lab-setting, four studies have a level of evidence 1b or $2b^{42, 47, 53, 54}$. These studies show construct validity, high internal consistency and variable inter-rater reliability for OSATS used with gynaecologic bench tasks. Other studies have a level of evidence 3b or 4, but show similar results. For use in the OR, no evidence of a high level is available

| Table | s specif | Ications al | nd most important results f | or all include | d studie | s addressing Obje | ective Structured A | vssessment c | of Technical Skill | s (USAIS). | |
|-------|----------|-------------|-----------------------------|--------------------|----------|--------------------------|--------------------------|-------------------|--------------------|-------------------|-------------------|
| Ref | Setting | Mode | Task/Procedure | Level | z | Constru | uct validity | | Reli | ability | |
| | | | | | | | | Internal c | onsistency | Inter-rater | reliability |
| | | | | | | GRS | Checklist | GRS | Checklist | GRS | Checklist |
| 53 | Lab | Video | Episiotomy repair | 1b/2b ^a | 40 | Yes | Yes | NA | NA | 0.59 | 0.80 |
| 50 | Lab | Live | | 4 | 18 | Yes | Yes | 0.95 | 0.95 | NA | NA |
| 42 | Lab | Live | Cystoscopy & | 1b/2b ^a | 55 | Yes | Yes | 0.93 ^d | 0.72 ^d | 0.69 ^d | 0.68 ^d |
| | | | colposuspension | | | | | 0.92 ^e | 0.85 [€] | 0.84 ^e | 0.92€ |
| 54 | Lab | Live | Hysteroscopy | $1b/2b^{a}$ | 48 | Yes | Yes | 0.98 | 0.79 | 0.42-0.93 | 0.75 |
| 47 | VRS | Live | Myoma resection | 2b | 13 | Yes | No | 0.98 | 0.70 | 0.88 ^h | 0.78 ^h |
| 45 | Lab | Live | Basic skills for | 4 | 24 | Yes, r>0.8 ^b | Yes, r >0.8 ^b | 0.89 | >0.8 | 0.87 | >0.8 |
| 44 | Lab | Live | gynaecological | 4 | 24 | Yes, r>0.8 ^b | Yes, $r > 0.8^{b}$ | 0.94 | 0.77 | 0.91 | 0.92 |
| 46 | Lab | Live | trainees | 3b | 16 | Yes | Yes | 0.96 | 0.96 | 0.95 ^h | >0.8 ^h |
| 43 | Lab | Live | | 3b | 116 | Yes | Yes | 0.97 | 0.95 | 0.95 ^h | 0.88 ^h |
| 49 | Lab | Live | Basic surgical | 3b | 20 | Yes | Yes | <0.8 | <0.8 | <0.8 | <0.8 |
| 51 | Lab | Live | skills | 3b | 48 | Yes | Yes | 0.84 | 0.78 | NA | NA |
| 38 | Lab | Live | | 4 | 77 | Yes | Yes | 0.82 | <0.8 | NA | NA |
| 41 | Lab | Live | | Зb | 9 | | r <0.8℃ | NA | NA | NA | NA |
| 61 | Lab | Live | | 3b | 53 | 0.70 | 0.58 | 0.85 | 0.79 | NA | NA |
| 52 | Lab | Video | Small bowel | 3b | 40 | NA | NA | NA | NA | >0.8 | NA |
| | | | anastomosis | | | | | | | | |
| 59 | Lab | Video | Vascular anastomosis / | 3b | 50 | Yes, r=0.59 ^f | No | NA | NA | 0.89 | 0.81 |
| 58 | Lab | Video | patch | 3b | 23 | Yes | Yes | NA | NA | | |
| 26 | Lab | Video | LAR and Nissen | 3b | 29 | No | NA | NA | NA | 0.72 | NA |
| 57 | ARS | Video | 3 tasks on ProMIS | 3b | 20 | r=0.77 ⁹ | r=0.77-0.81 ⁹ | NA | NA | 0.93 | 0.88 |
| 40 | OR | Live | Different gynaecologic | 3b | 119 | Yes ⁱ | Yes ⁱ | NA | NA | NA | NA |
| 62 | OR | Live | FC | 3b | 22 | Yes | NA | NA | NA | 0.28 | NA |
| 09 | OR | Video | | 3b | 22 | No | NA | NA | NA | 0.57 | NA |
| 56 | OR | Video | ГC | 3b | 47 | Yes | NA | 0.72 ^k | NA | 0.72 | NA |
| | | | | | | | | | | | |

Objective assessment of technical surgical skills

CHAPTER 2.1

| | • | Mode | Task/Procedure | Level | z | Constru | uct validity | | Re | liability | |
|-----|---------------|----------------|--------------------------------|---------------|-----------|--|--|----------|-------------|-----------|----------------|
| | | | | | | | | Internal | consistency | Inter-rat | er reliability |
| | | | | | | GRS | Checklist | GRS | Checklist | GRS | Checklist |
| | OR | Video | Salpingectomy | Зb | 21 | Yes | Yes | NA | NA | 0.83 | 0.83 |
| 5 | OR | Live | 3 different operations | 4 | 41 | Yes r=0.89⁵ | Yes r=0.89⁵ | NA | NA | 0.73 | 0.78 |
| 6 | OR | Live | Carotid endarteriectomy | 4 | 28 | r=0.69-0.82 ^g Yes ^j | r=0.21-0.83 ^g Yes ⁱ | NA | NA | NA | NA |
| ٩A | Not a | ddressed | | | | | | | | | |
| GRS | Globa | al rating sca | le | | | | | | | | |
| ARS | Augr | nented reali | ty simulator | | | | | | | | |
| /RS | Virtua | al reality sim | nulator | | | | | | | | |
| AR- | Low-ĉ | anterior rese | ection | | | | | | | | |
| _ | 1b foi | r GRS, 2b fo | r checklist | | | | | | | | |
| 0 | Conct | urrent valid | ity: correlation of GRS with u | checklist | | | | | | | |
| | Conci | urrent valid | ity: correlation of OSATS scc | ore with facu | Ity ratin | Sc | | | | | |
| - | For cy | /stoscopy | | | | | | | | | |
| | For cc | alposuspen: | sion | | | | | | | | |
| | Conct | urrent valid | ity: correlation of OSATS GR | S score with | ICSAD n | notion analysis | | | | | |
| Ē | Const | truct validity | y: correlation with experiend | ce | | | | | | | |
| ~ | Corre | lation betw | een blinded and un-blinder | d raters | | | | | | | |
| | Face v | validity | | | | | | | | | |
| | Discri | iminant vali | dity | | | | | | | | |
| | Inter- | test reliabili | ity | | | | | | | | |
| | Regre | ssion coeffi | icient for relation with postg | graduate yea | ır | | | | | | |

as all 7 studies offer level 3b or 4 evidence. These 7 studies show construct validity and sporadically address reliability, which is only above 0.8 in one study⁴⁸.

Motion analysis

Motion analysis uses parameters that are extracted from motion of the hands or laparoscopic instruments. Nineteen studies were identified concerning this method of assessment^{25, 56-59, 63-75}. These studies addressed 5 different instruments: the Imperial College Surgical Assessment Device (ICSAD; Imperial College, London, UK), the Advanced Dundee Psychomotor Tester (ADEPT; University of Dundee, Dundee, Scotland), the ProMIS[™] Augmented Reality Simulator (Haptica, Dublin, Ireland), the Hiroshima University Endoscopic Surgical Assessment Device (HUESAD; Hiroshima University, Hiroshima, Japan) and the TrEndo Tracking System (Delft University of Technology, Delft, Netherlands). (See Table 4 for details.)

Nine studies addressed the (ICSAD)^{25, 56, 58, 59, 63-65, 73, 75}. Construct validity was established, mostly for the parameters time and for number of movements. Only Aggarwal⁵⁶ used the ICSAD in the OR. Inter-test reliability was not found to be high and only moderate correlation existed with OSATS, which was considered the current gold standard for objective assessment. In study by Datta⁵⁹, the same correlation with OSATS was found, although it was used in a lab-setting and not in the OR. Levels of evidence of all studies were consistent level 3b.

The ADEPT showed construct validity for one out of three parameters in a level 1b study by Francis⁶⁸. Two other studies addressed correlation with clinical assessment and reliability, but these have a lower level of evidence and fewer participants^{67, 69}.

ProMIS[™] is a hybrid simulator, which combines a live and virtual environment. Tasks on this simulator are done in a box-trainer, but a virtual interface is placed over the image of the camera in the box trainer. Two other cameras are used for motion tracking of the instruments. In a level 2b study by VanSickle⁷², construct validity was established and internal consistency was 0.95. However, this study only used 10 participants. Other studies used more participants, but have lower levels of evidence (see Table 4).

The HUESAD was developed to analyze movements in vertical and horizontal planes. In a study by Egi⁶⁶ construct validity was established, comparing novices and experts. However, this was the only study about the HUESAD, and it offers level 3b evidence.

Our group designed the TrEndo tracking system for motion analysis, to be used in a box trainer. In a study by Chmarra⁷⁶ participants were classified as novice, intermediate or expert by analysis of 6 motion analysis parameters (time, depth perception, path length, motion smoothness, angular area and volume). The data of these 6 parameters was first compressed using Principal Component Analysis (PCA) and subsequently classified using Linear Discriminant Analysis (LDA). In this way, 23 out of 31 participants were correctly classified. Further research with the TrEndo is in progress.

| Ref | Setting | Device | Level | N | (| Construct valid | ity | Other validity | Reliability |
|-----|---------|---------|-------|-----|------------------------|------------------------|------------------------|------------------------|--------------------------------------|
| | | | | | Time | No of movements | Path length | | |
| 56 | OR | ICSAD | 3b | 47 | Yes ^a | Yesª | Yesª | <0.63° | 0.63 ^b |
| 65 | Lab | | 3b | 51 | 0.66 ^d | 0.76 ^d | No | | |
| 59 | Lab | | 3b | 50 | Yes | Yes | NA | < 0.60° | |
| 75 | Lab | | 3b | 30 | NA | NA | NA | 0.49-0.65 ⁱ | |
| 64 | Lab | | 3b | 30 | Yes | Yes | NA | | |
| 58 | Lab | | 3b | 23 | Yes | Yes | NA | | |
| 63 | Lab | | 3b | 30 | Yes | Yes | NA | | |
| 73 | Lab | | 3b | 15 | Yes | Yes | Yes | 0.78 ^e | |
| 25 | Lab | | 3b | 26 | Yes | NA | Yes | | |
| | | | | | Time | Error score | Task score | | |
| 69 | Device | ADEPT | 3b | 10 | NA | NA | NA | 0.789 ^f | |
| 67 | Device | | 3b | 20 | NA | NA | NA | | 0.6 ^b ; 0.98 ^g |
| 68 | Device | | 1b | 40 | No | Yes | No | | |
| | | | | | Time | Smoothness | Path length | | |
| 72 | Device | ProMIS™ | 2b | 10 | Yes | Yes | Yes | | 0.95 ⁹ |
| 70 | Device | | 3b | 30 | 0.78 ^d | 0.75 ^d | 0.67 ^d | 0.18-0.88 ^h | |
| 71 | Device | | 3b | 160 | Yes | Yes | Yes | | |
| 57 | Device | | 3b | 20 | 0.61-0.81 ^d | 0.36-0.98 ^d | NA | | |
| 74 | Device | | 3b | 46 | 0.07-0.60 ^d | 0.11-0.59 ^d | 0.00-0.39 ^d | | |
| 66 | Lab | HUESAD | 3b | 37 | Yes | | | NA | NA |
| 76 | Lab | TrEndo | 3b | 31 | Yes ^j | | | NA | NA |

| Table 4 | Specifications and most im | portant results for all included | studies addressing | g motion analy | ysis |
|---------|----------------------------|----------------------------------|--------------------|----------------|------|
|---------|----------------------------|----------------------------------|--------------------|----------------|------|

NA Not addressed

a Only for dissection part of LC

b Inter-test reliability

- c Concurrent validity: correlation with OSATS
- d Correlation with experience
- e Correlation of path length with procedure-specific checklist
- f Concurrent validity: correlation overall performance on ADEPT with clinical assessment
- g Internal consistency
- h Correlation with GRS
- i Correlation of 'surgical efficiency score', based on ICSAD motion analysis, with OSATS

j Classified 74% of participants correctly using Linear Discriminant Analysis of motion analysis parameters

Virtual reality simulators

Virtual reality simulators are especially known as trainers for endoscopic motor skills. As several parameters of performance are measured, VR simulators may be used for assessment of skills as well. Twenty-six studies were identified which addressed this aspect of VR simulators⁷⁷⁻¹⁰². Levels of evidence ranged from 1b to 4. (See Table 5 for details.)

Studies on five different simulators were identified: Minimally Invasive Surgical Trainer Virtual Reality (MIST[™] VR; Mentice, Göthenburg, Sweden), LapSim (Surgical Science, Göthenburg, Sweden), LAP Mentor[™] (Simbionix Corporation, Cleveland, Ohio, USA),

Xitact[®] LS 500 (Mentice, Göthenburg, Sweden) and Simulator for Endoscopy SIMENDO[®] (DelltaTech, Rotterdam, the Netherlands). These simulators all provide tasks to train basic surgical skills for general surgery, gynaecology or laparoscopy in general. For assessment, most simulators use simple motion analysis parameters, like path length or economy of motion and all use time to task completion. Some use a composite score, which is a simple sum or predetermined by the manufacturer and different for every task, while others use error scores.

For all five simulators most studies show good results. However, most studies offer level 3b evidence and therefore the results should be considered carefully. Studies with higher levels of evidence exist for MIST[™] VR, LAP Mentor[™] and LapSim.

Two studies by Gallagher^{81, 82} with level 1b evidence and one level 2b study by Taffinder⁹⁴ established construct validity for MIST[™] VR parameters. Furthermore, in a study by Aggarwal¹⁰³ the use of proficiency scores was stated. Other studies found comparable results, but have a lower level of evidence^{89, 94}. In a study by Cope⁷⁸, on the other hand, construct validity could not be established for any parameter and in two studies by Grantcharov⁸³ and Madan⁸⁷ poor correlation with performance on a pig was found (i.e., concurrent validity). These studies offer level 3b evidence as well.

For LAP Mentor[™], a study by Zhang consisted of level 1b evidence and showed construct validity for time and composite score comparing novices and residents⁹⁹. In a level 3b study by Aggarwal construct validity was established for most tasks, by different parameters. Moreover, in this study cut-off values were defined¹⁰¹. In other studies results for LAP Mentor[™] were less consistent (see Table 5).

LapSim has been studied extensively and most studies show construct validity. One study, by Aggarwal¹⁰², offers level 1b evidence and shows construct validity for time and path length for all exercises. Cut-off values were defined in this study and in a study by Sherman as well⁹³.

Video assessment

A separate category was defined for video assessment. With video assessment, a task, performance or operation is video taped and is rated at a later point in time, which adds to its flexibility. The methods for assessment are the same as in live settings, but the fact that the performance is video taped may have a considerable impact on the outcome of the assessment. For example, often only the laparoscopic camera shot is taped and not the whole OR, which may blind the observer to certain aspects of the operation.

Five studies were identified that explicitly addressed the impact of the use of video registration on the outcome of the assessment^{60, 75, 104-106}. (See Table 6 for details.)

Studies by Beard^{104, 105} and Driscoll¹⁰⁶ established construct validity for video assessment, with level 1b in one study by Beard¹⁰⁴. However, in this study only 2 videos of 2 subjects with a large difference in performance level (inexperienced vs. experienced)

| Ref | Simulator | Level | Ν | | | Construct vali | dity | |
|-----|----------------------------|-------|-----|--------------------|------------------------|-------------------------|------------------------|------------------------|
| | | | | Time | Economy of movement | Economy of diathermy | Errors | Composite score |
| 81 | MIST [™] VR | 1b | 36 | Yes | Yes | No | Yesª | NA |
| 82 | | 1b | 36 | Yes | Yes | Yes | Yes | NA |
| 94 | | 2b | 30 | Yes | Yes | NA | Yes | NA |
| 78 | | 3b | 22 | No | No | NA | No | NA |
| 83 | | 3b | 14 | NA | NA | NA | 0.5-0.7 ^c | 0.05-0.80 ^c |
| 89 | | 3b | 8 | Yes | Yes | NA | Yes | NA |
| 87 | | 3b | 32 | < 0.56° | < 0.56° | NA | NA | 0.21-0.56 ^c |
| | | | | Time | Path length | Angular path | Errors | Composite score |
| 80 | LapSim | 3b | 24 | Yes | Yes | Yes | Yes | NA |
| 85 | | 3b | 115 | Yes | Yes | Yes | Yes | NA |
| 86 | | 3b | 32 | Yes | Yes | Yes | Yes ^d | Yes |
| 79 | | 3b | 54 | Yes | Yes ⁿ | NA | Yes ^d | Yes |
| 93 | | 3b | 24 | Yes ^e | Yes ^f | NA | NA | NA |
| 84 | | 3b | 10 | 0.74 ^g | 0.69-0.98 ⁹ | NA | 0.67-0.89 ⁹ | NA |
| 97 | | 3b | 34 | Yes | Yes | Yes | Yes | NA |
| 100 | | 4 | 47 | <0.51 ^h | NA | NA | 0.01-0.42 ^h | NA |
| 102 | | 1b | 40 | Yes | Yes | NA | No | NA |
| | | | | Time | No of movements | Economy of movement | Speed | Composite score |
| 99 | LAP Mentor™ | 1b | 27 | Yes | NA | NA | NA | Yes |
| 98 | | 3b | 31 | Yes | Yes ⁱ | Yes ⁱ | Yes ⁱ | NA |
| 88 | | 3b | 103 | NA | NA | NA | NA | Yes ⁱ |
| 77 | | 3b | 27 | Yes ^k | NA | NA | NA | Yes ^k |
| 101 | | 3b | 57 | Yes° | Yes ^d | Yes ^d | Yes ^p | NA |
| | | | | Time | Path length | Economy of movement | Speed | Composite score |
| 91 | Xitact [®] LS 500 | 3b | 20 | No | Yes | NA | No | NA |
| 90 | | 3b | 307 | Yes | Yes ⁱ | NA | NA | NA |
| 92 | | 3b | 74 | Yes ^m | NA | NA | NA | Yes |
| | | | | Time | Path length | Errors | | |
| 96 | SIMENDO® | 3b | 25 | Yes | Yes ⁱ | Yes | | |
| 95 | | 3b | 61 | Yes | Yes | No | | |

 Table 5
 Specifications and most important results for all included studies addressing virtual reality simulators.

- NA Not addressed
- a Only significant difference between experts and intermediates
- b Internal consistency
- c Concurrent validity: correlation with performance on pig
- d Only for 1 task
- e Construct validity for 'summary measure' time-error
- f Construct validity for 'summary measure' motion economy, with 2 of 3 tasks
- g Predictive validity: correlation with performance in the OR
- h Concurrent validity: correlation with performance on box trainer
- i Only for non-dominant hand
- j Only for 1 of 8 tasks
kOnly for 2 of 6 tasksIOnly for right instrumentmAlso concurrent validity: subject with max score on pelvitrainer had significantly shorter task
time on XitactnOnly for 2 of 5 tasksoFor 6 of 9 skillspFor 4 of 9 skills

were rated by different groups of raters. In the other study by Beard, a good correlation was found between video and live assessment, although that study offers level 3b evidence¹⁰⁵.

Editing of video tapes alters the assessment. In the studies by Beard and Driscoll, raters were permitted to fast forward the tape at their own discretion. In two studies by Scott⁶⁰ and Datta⁷⁵ the effect of editing video tapes before rating them was studied. In the study by Scott⁶⁰ the video tapes were shortened to 10 minutes, showing only the essential parts. A poor correlation with live assessment was found. In the study by

| Ref | Setti | ing Procedure | Level | Ν | Const valie | truct dity | Correlation video-live | | Relia | ability | |
|---|-------|---|--|---|--|-------------------------|------------------------------|--------------------|------------|---------------------|-------------------|
| | | | | | | | | Interno consist | al ency | Inter-ro reliabi | ater lity |
| | | | | | Video | Live | | Video | Live | Video | Live |
| 60 | OR | LCª | 3b | 22 | No | Yes | <0.33 | NA | NA | 0.28 | 0.57 |
| 105 | OR | SFD ^b | 3b | 33 | NA | NA | 083-0.92 | NA | NA | NA | 0.91 ^e |
| 106 | OR | Inguinal hernia repair ^{b,c} | 3b/4 ^h | 9 | Yes | No | NA | >0.76 | >0.85 | >0.69 | NA |
| 75 | Lab | Vascular & bowel anastomosis ^d | 3b | 30 | <0.37 ^f | NA | NA | NA | NA | 0.59 -0.80 | NA |
| 104 | OR | SFD [♭] | 1b | 2 ^g | Yes | NA | NA | NA | NA | NA | NA |
| LC SFD a b c d e f | | Laparoscopic chole Saphenofemoral di Edited video tapes: Fast forwarding of Edited video tapes: Edited video tapes: Inter-test reliability Concurrent validity | cystecto sconnec length video ta only es length | omy ction 10 mi pe wa sentia 2 min tion o | nutes as permitt Il steps sh utes of full lend | ted Iown ath vide | o score with | snapsho | t video | score | |
| g h (leve | l 4) | Two video tapes were b | ere shov linded (| vn to level 3 | 14 surgeo 3b), real-t | ons, 14 t ime ass | rainees and 1 essment was | 3 OR nu not | rses | Score | |

Table 6Specifications and most important results for all included studies addressing effect of videotaping.

Datta⁷⁵, a 2 minute snapshot tape of a task was recorded and the rating was compared with the rating of a full length video tape. The results were comparable to those of Scott.

Miscellaneous

Nine studies did not fully fit any other category^{75, 107-114}. Levels of evidence of these studies extended from 2b to 4 and concerned six different methods for assessment. (See Table 7 for details).

One method of specific interest is outcome measurement, as it is often applied in clinical practice. With this method, numbers of complications, morbidity and mortality are kept in logbooks or portfolios. Haddad¹⁰⁷ compared complications between junior and senior surgeons in a study with 691 cases, which were separated for the extent of the operation. A difference was found for moderate extensive operations, for which more complications were attributed to senior surgeons. However, this difference was considered to exist due to allocation of more difficult cases to more senior surgeons, leading to a bias. Therefore, patient outcome was not considered to be a useful method for assessment.

| Ref | Setting | Mode | Method | Level | Ν | Validity | Relia | bility |
|-----|---------|-------------------|--|-------|-----|------------------------|-------------------------|----------------------------|
| | | | | | | | Internal consistency | Inter-rater reliability |
| 110 | Lab | Live | Force/torque metrics | 2b | 4 | Yes ^a | NA | NA |
| 111 | Lab | Live | (Hidden) Markov | 2b | 8 | Yes ^a | NA | NA |
| 109 | Lab | Live | modeling | 2b | 10 | Yes ^a | NA | NA |
| 108 | Lab | Live | | 3b | 11 | 0.93 ^b | NA | NA |
| 75 | Lab | Video | Quality of final product | 3b | 30 | 0.34-0.55 ^b | NA | 0.80-0.84 |
| 113 | Lab | Video | Time, errors and needle manipulations for suturing | 3b | 32 | Yes ^{a, f} | NA | 0.86-0.91 ^g |
| 107 | OR | Data ^e | Patient outcome | 4 | 691 | No | NA | NA |
| 114 | OR | Live | | 4 | 29 | Yes ^c | NA | NA |
| 112 | Lab | Video | Error scoring | 4 | 60 | >0.8 ^d | NA | NA |

| Table 7 | Specifications and most important results for all included studies addressing miscellaneous |
|---------|---|
| methods | for assessment. |

NA Not addressed

a Construct validity

b Concurrent validity: correlation with OSATS

c Leakage in lab-task for vascular anastomosis predicted leakage in OR and time in OR

d Concurrent validity: correlation with Objective Structured Clinical Examination (OSCE) for performance in simulated LC

e Prospective data collection

f Good for time, limited for errors and needle manipulations

g Tapes were rescored until reliability was above 0.80

Another interesting method is the use of (hidden) Markov modeling. This is a mathematical way of compressing large amounts of data and producing one single measure to indicate a subject's distance from an ideal learning curve. With the studies shown in Table 8, this method was used to compress data of motion and force/torque. These three studies, of which two consisted of level 2b evidence, established construct and concurrent validity^{108, 109, 111}. However, in these studies very limited numbers of participants were used, so larger studies are needed to provide more solid evidence to show whether this method can truly distinguish between individuals with different performance.

Fundamentals of Laparoscopic Skills (FLS)

Assessment of the FLS manual skills test is based on the McGill Inanimate System for Training and Evaluation of Laparoscopic Skills (MISTELS). It consists of 5 tasks which are rated by two metrics: 'time to complete the task' and 'accuracy', calculated by predetermined penalties. As FLS consists of this fixed set of box trainer tasks which is used for assessment and is not a specific rating method, it would actually not fit the inclusion

| Ref | Level | Ν | Val | lidity | Reliability | Sensitivi | ty/Specificity |
|-----|-------|-----|---------------------------------------|------------------------|--|------------------|------------------|
| | | | Construct | Concurrent | | Sensitivity | Specificity |
| 117 | 1b | 50 | NA | 0.51ª | NA | NA | NA |
| 122 | 3b | 12 | NA | NA | 0.77-0.86 ^b 0.98-1.00 ^c 0.37-0.89 ^d | NA | NA |
| 116 | 4 | 42 | 0.26-0.69 ^e | NA | NA | NA | NA |
| 115 | 4 | 10 | Yes ^f 0.82 ^e | NA | NA | NA | NA |
| 118 | 4 | 165 | Yes | NA | NA | 82% ^g | 82% ^g |
| 119 | 4 | 12 | NA | 0.15-0.76 ^h | NA | NA | NA |
| 120 | 4 | 215 | Yes | 0.81 ^h | NA | NA | NA |
| 121 | 4 | 58 | Yes | NA | NA | NA | NA |
| 36 | 4 | 40 | Yes | 0.77 ⁱ | NA | 91% ^j | 86% ^j |

Table 8Specifications and most important results for all included studies addressing Fundamentals of
Laparoscopic Skills (FLS) manual skills test.

| 36 | 4 | 40 | Yes | 0.77 ⁱ | NA | 91% ^j | 86% ^j |
|----|---|------------------|---------------|-------------------|--------------|------------------|------------------|
| NA | | Not addresse | d | | | | |
| а | | Correlation w | ith In-Train | ing Evaluation Re | ports (ITER) | | |
| b | | Internal cons | istency | | | | |
| с | | Inter-rater rel | iability | | | | |
| d | | Inter-test relia | ability | | | | |
| e | | Correlation o | f total score | with postgradua | ate year | | |
| f | | For 2 of 3 tasl | KS | | | | |
| g | | For a total sco | ore cut-off o | of 270 | | | |
| h | | Correlation w | ith perform | nance in vivo | | | |
| i | | Correlation w | ith GOALS- | score | | | |
| j | | For a mean so | core cut-off | of 70 | | | |

criteria for this review (see methods section). However, FLS is the official manual skills test for surgical residents in the USA and is also used in other countries. Therefore, it is considered very useful to include in this review.

Nine studies were identified concerning MISTELS^{36, 115-122}. (See Table 8 for details.) Construct validity was established in six studies^{36, 115, 116, 118, 120, 121}, of which four studies^{36, 118, 120, 121} found highly significant differences between subjects with different training levels, and two studies^{115, 116} found correlations with training levels varying from poor to good. Two large studies were produced by Fraser¹¹⁸ and Fried¹²⁰ with 165 and 215 subjects respectively. Unfortunately, these studies only offer level 4 evidence, because raters were not blinded to the training level of the participants.

Four studies addressed concurrent validity, comparing MISTELS with other assessment methods. One study offers level 1b evidence and found a moderate correlation with In-Training Reports (ITER)¹¹⁷.

A study by Vassiliou¹²² is the only study in which reliability of MISTELS was studied. Good internal consistency and excellent inter-rater reliability were found, tested by comparing blinded to unblinded raters. This is the only study that clearly stated the use of blinded raters. It consists of level 3b evidence as it used 12 non-consecutive participants.

Finally, in 2 studies a cut-off value was calculated for the FLS-score to use for certification^{36, 118}. Such a cut-off value is essential for summative assessment. Both studies offer level 4 evidence, since the raters were not blinded.

DISCUSSION

This study provides an overview and qualification of current methods for objective assessment of technical surgical skills in order to form a guideline for their use in clinical practice. As stated in the introduction, methods for objective assessment are needed for assessing trainees' performance, but also assessing performance of practicing surgeons. In an era with focus on new insights in training and on quality and safety of surgery, these are important issues with potentially major impact. This requires solid proof of validity and reliability of assessment methods.

From all studies included in this review it can be concluded that OSATS is probably most accepted as the gold standard for objective skills assessment. However, a high level of evidence for OSATS is only reached for use with gynaecologic bench tasks in a lab-setting. Evidence for use in the OR is of lower grade and less abundant. Therefore, it is doubtful whether OSATS can distinguish between different levels of performance in the OR. Studies by Martin⁴⁹ and Beard¹⁰⁵ were the only studies to correlate performance in bench tasks with performance in live animals (pigs) and the OR, finding moderate

correlations, not exceeding 0.8. Furthermore, cut-off values have not been defined for OSATS. These shortcomings should not prevent its use for feedback and discussion (formative assessment), but do prohibit the making of more important decisions based on OSATS. Therefore, although it might be the current gold standard, OSATS should not be used for summative assessment in the OR if these results are to be used for important decisions other than simple evaluation.

The same shortcomings affect other methods of assessment. Procedure-specific checklists and global rating scales (GRS) have mostly been studied in single studies. Only for one procedure-specific checklist a study with a high level of evidence is available and shows good results for its use with LC²¹. GOALS is the only GRS which has been studied in multiple studies, but only one of these studies offers a high level of evidence³⁷. Therefore, the evidence is limited for both the mentioned checklist and the GOALS. Their use for formative assessment is an option, but using them for summative assessment is not recommended.

The use of motion analysis devices and virtual reality simulators for skills assessment has been studied more extensively. For motion analysis, construct validity was established for the ICSAD and the ProMIS[™] in multiple studies. Therefore, it seems plausible that these devices can differentiate between levels of performance and can be used for formative assessment. However, all these studies offer a low level of evidence and the only study for the ProMIS[™] which offers level 2b evidence, used 10 participants⁷². Moreover, cut-off values for the scores have not been defined. Consequently, these devices should only be used for formative assessment. For other motion analysis devices, fewer studies are available, meaning there is not enough evidence to use them for summative assessment.

Virtual reality simulators have all been tested in multiple studies and all simulators seem to be able to distinguish in performance in some way. However, high level evidence is scarce and only available for MIST[™] VR, LAP Mentor[™] and LapSim. The results for MIST[™] VR are consistent and, together with proficiency scores, form the evidence for its use for summative assessment. The results for LAP Mentor[™] are not. The only study with a high level of evidence shows construct validity for 2 parameters. However, cut-off values have been defined for this simulator. Therefore, summative assessment is possible, but not recommended. For LapSim, all studies, including one level 1b study, show construct validity and cut-off values have been defined. Therefore, the evidence is solid enough to use LapSim for summative assessment.

The results for video assessment show that it is possible to use video tapes to distinguish between individuals with a large difference in performance levels. No evidence exists for distinguishing between more subtle differences in performance level. Consequently, using video tapes for summative assessment is not yet possible. Moreover, it is shown that editing, especially shortening videos, before assessment has a clear impact on the outcome of assessment. This hampers the practical applicability of video assessment on a large scale.

The FLS manual skills test was broadly studied and results were mostly good. A main shortcoming is that only one study used blinded raters¹²². Therefore, all other studies have lower levels of evidence. However, the impact of blinding on scoring of the MIS-TELS tasks is questionable, because the predefined errors can not easily be interpreted in different ways and can therefore be indicated as objective measures^{115, 116, 118}. Likewise, the time to complete a task is an objective measure. Moreover, in a study by Vassiliou¹²² a very high correlation was found between blinded and unblinded raters. Consequently, one could argue that the results in all studies without blinding are unlikely to change much when raters would be blinded and these studies could be considered as having a higher level of evidence. Therefore, FLS seems well suited for formative assessment, and as cut-off values have also been defined, it could even be used for summative assessment.

Certain factors made it difficult to compare different methods for assessment. Much heterogeneity exists among the included studies, because existing methods for assessment have been adapted for use in other studies, and methodology and statistics differ. Different settings or tasks were used. Variation in outcomes of these studies could therefore be caused by these differences instead of shortcomings of the assessment methods.

To determine the quality of the studies included in this review, the Oxford Centre for Evidence-based Medicine Levels of Evidence for 'diagnostic studies' were used. This category provided the best fit with the design of validating studies. For this category, high levels of evidence are only reached when blinded reference standards (i.e., the observer is blind to the training level of the subject) are used and when cohorts of study objects consist of consecutive participants. Unblinded rating is less objective, which affects not only reliability, but validity as well, as the validity relies on the outcome of a test which in turn is influenced by reliability. Non-consecutive cohorts carry the risk of selection bias because certain eligible subjects might have been deliberately excluded. In many articles it was not clearly stated whether the cohorts were consecutive. These studies were qualified as level of evidence 3b. This explains the high number of studies with this level of evidence. One might argue that this approach is not completely fair. It is common with this kind of studies that all eligible subjects are asked to voluntarily participate in a study, which might result in a consecutive cohort. This fact was certainly considered, but as it was explicitly mentioned in a number of studies that all eligible subjects were included, it was considered appropriate to stick to the strict definition of the Oxford Levels of Evidence. Moreover, it is believed by the authors that customizing these levels would inflict on its familiarity.

For reliability a threshold of 0.8 was chosen for 'good' reliability. As explained in the methods section this choice was based on the literature. However, one could criticize

that this threshold is too rigid and not demanded for all assessments as they not all need to be that strict. While that may be true, this threshold was adhered to, to be able to distinguish methods that are sufficiently reliable for high-stake assessment. Moreover, many studies show reliability for their method of assessment above this threshold, indicating that it is possible to meet the requirement.

As already slightly mentioned above, studies on assessment methods are hampered by the fact that it is difficult to set up a good study with participants who work in clinical practice. Time is always short and schedules not very flexible, which makes it hard to compose large and consecutive groups of participants. Using blinded raters is even a bigger challenge. As they have to be blinded to the assessed subject's training and clinical level, they have to originate from another hospital at least, which brings along complex logistic organization. Setting up more solid studies, which meet important criteria, requires that these issues be solved. Although enormous progress has been made, solid studies are required to be able to draw firm conclusions on the matter of skills assessment. These studies should adhere to criteria which ensure true measures of validity and reliability, as mentioned above and addressed by the levels of evidence. One solution to facilitate this might be to embed future studies within residency programs or training curricula. In this way, participation of all eligible subjects and use of blinded raters will be easier to realize. Consequently, this would require cooperation of residency program directors.

Different methods for skills assessment are available and are appropriate in different situations. Which assessment method will be most appropriate depends on different aspects. The first issue to consider is the type of assessment and its consequences. To enable formative assessment, a method will have to give a reasonable impression of a subject's performance, as this form of assessment is only used as input for feedback and tracking a subject's progress over time. A negative outcome will not have direct consequences in terms of career progression or certification. On the contrary, such consequences might result, although not necessarily, from summative assessment. Therefore, if methods are to be used for this form of assessment, these should be highly accurate and reliable, as they might be used to test a subject's performance against predefined criteria. A crucial element for this is defining cut-off values for the scores which serve as these criteria. Furthermore, it should be clear what kind of performance has to be assessed. It has to be considered whether it is general skills which are of interest, or precise and systematic completion of a specific procedure. The latter requires a procedure-specific checklist, although this might not always give a proper representation of skills in general. In addition, it is important to realize that the value of a good assessment method can diminish when it is used in an inappropriate setting. Different methods are validated for different settings, which simply limits the applicability of a certain method. An assessment method which is suitable for standardized tasks in a lab-setting may not be suitable for a real operation in the operating room, where performance and outcome are easily influenced by many factors. Formulating clear answers to the issues discussed above is of great importance and should be adhered to when selecting a proper method for assessment of skills.

In general, the field of skills assessment has moved a great deal forward and an important change of environment has started to take place. Traditionally, assessment was subjective or even anecdotal. At present, many objective assessment methods of skills are available and progress is continuously being made. However, caution is warranted in applying assessment methods in practice. Methods have been studied more extensively in the lab-setting than in the operating theatre. The quality of the studies is mostly good enough to justify formative assessment, but still too low for summative assessment. The latter remains an important challenge. As the pioneering work has been done, future studies should be well structured to solve remaining issues. Methods suitable for summative assessment would enable more properly selecting candidates for surgical training. Authorities await the moment such methods can be applied for quality assurance of surgical performance, but this point has not yet been reached. Most importantly, it should be kept in mind that technical skills are just one of the competencies of a surgeon and that there are many other qualifications that make a good surgeon.

CONCLUSION

Important changes have moved the field of skills assessment in a more objective direction and this is still ongoing. Based on the studies included in this review, it can be concluded that many different methods for objective assessment of technical skills currently exist and have been studied abundantly. However, for most of these methods the available evidence was considered not to be of the appropriate quality to justify their unconditional use for summative assessment, such as examination or even credentialing. On the contrary, using these methods for feedback is less demanding and most methods are well suited for this purpose.

Before choosing a method for assessment, it is essential to know how, where, and for what purpose that method will be used. This knowledge will help to prevent a method from being used beyond its purpose.

REFERENCES

- 1. Reznick RK. Teaching and testing technical skills. *Am J Surg* 1993;**165**(3): 358-361.
- 2. Halsted WS. The training of the surgeon. *Bulletin of the Johns Hopkins Hospital* 1904;**15**: 267-275.
- Darzi A, Smith S, Taffinder N. Assessing operative skill. Needs to become more objective. BMJ 1999;318(7188): 887-888.
- 4. Risico's minimaal invasieve chirurgie onderschat. http://www.igz.nl/15451/475693/2007-11_Rapport_Minimaal_in1.pdf.
- Darzi A, Datta V, Mackay S. The challenge of objective assessment of surgical skill. Am J Surg 2001;181(6): 484-486.
- 6. Kohn LT CJ, Donaldson MS (ed). *To err is human: building a safer health system*. National Academy Press: Washington DC, 1999.
- 7. Kopta JA. Approach to evaluation of operative skills. *Surgery* 1971;**70**(2): 297-&.
- Figert PL, Park AE, Witzke DB, Schwartz RW. Transfer of training in acquiring laparoscopic skills. J Am Coll Surg 2001;193(5): 533-537.
- Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. BMJ 2003;327(7422): 1032-1037.
- 10. Stassen LP, Bemelman WA, Meijerink J. Risks of minimally invasive surgery underestimated: a report of the Dutch Health Care Inspectorate. *Surg Endosc* 2009.
- 11. Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. *Br J Surg* 2004;**91**(12): 1549-1558.
- 12. Fried GM, Feldman LS. Objective assessment of technical performance. *World J Surg* 2008;**32**(2): 156-160.
- 13. Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Assessment of technical surgical skills. *Eur J Surg* 2002;**168**(3): 139-144.
- 14. Jaffer A, Bednarz B, Challacombe B, Sriprasad S. The assessment of surgical competency in the UK. *Int J Surg* 2009;**7**(1): 12-15.
- 15. Phillips B BC, Sackett D, Badenoch D, Straus S, Haynes B, Dawes M. Oxford Centre for Evidencebased Medicine Levels of Evidence. 2001.
- Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surg Endosc* 2003;**17**(10): 1525-1529.
- 17. Carter FJ, Schijven MP, Aggarwal R, Grantcharov T, Francis NK, Hanna GB, et al. Consensus guidelines for validation of virtual reality surgical simulators. *Surg Endosc* 2005;**19**(12): 1523-1532.
- 18. Cooper. Individual Differences. Arnold: London, 1998.
- 19. 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;**189**(6): 566-574.
- 20. Larson JL, Williams RG, Ketchum J, Boehler ML, Dunnington GL. Feasibility, reliability and validity of an operative performance rating system for evaluating surgery residents. *Surgery* 2005;**138**(4): 640-647; discussion 647-649.
- 21. Sarker SK, Chang A, Vincent C. Technical and technological skills assessment in laparoscopic surgery. *JSLS* 2006;**10**(3): 284-292.
- 22. Sarker SK, Chang A, Vincent C, Darzi SA. Development of assessing generic and specific technical skills in laparoscopic surgery. *Am J Surg* 2006;**191**(2): 238-244.
- 23. Swift SE, Carter JF. Institution and validation of an observed structured assessment of technical skills (OSATS) for obstetrics and gynecology residents and faculty. In: 68th Annual Meeting of the

South-Atlantic-Association-of-Obstetricians-and-Gynecologists; 2006 Jan 29-Feb 01; Orlando, FL; 2006. p. 617-621.

- 24. Aggarwal R, Grantcharov T, Moorthy K, Milland T, Darzi A. Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Ann Surg* 2008;**247**(2): 372-379.
- Moorthy K, Munz Y, Dosis A, Bello F, Chang A, Darzi A. Bimodal assessment of laparoscopic suturing skills: construct and concurrent validity. *Surg Endosc* 2004;**18**(11): 1608-1612.
- 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. *Surg Endosc* 2004;**18**(12): 1800-1804.
- 27. Beckmann CRB, Lipscomb GH, Ling FW, Beckmann CA, Johnson H, Barton L. Computer-assissted video evaluation of surgical skills. *Obstet Gynecol* 1995;**85**(6): 1039-1041.
- Bramson R, Sadoski M, Sanders CW, van Walsum K, Wiprud R. A reliable and valid instrument to assess competency in basic surgical skills in second-year medical students. *South Med J* 2007;**100**(10): 985-990.
- 29. Chou B, Bowen CW, Handa VL. Evaluating the competency of gynecology residents in the operating room: validation of a new assessment tool. *Am J Obstet Gynecol* 2008;**199**(5): 571 e571-575.
- Delaney PV, Hennessy TP, Quill RD, Kalizser M. Assessment of operative surgical skills. *Ir Med J* 1978;**71**(13): 438-441.
- Doyle JD, Webber EM, Sidhu RS. A universal global rating scale for the evaluation of technical skills in the operating room. In: 93rd Annual Meeting of the North-Pacific-Surgical-Association; 2006 Nov 10-11; Spokane, WA; 2006. p. 551-555.
- Fung KFK, Fung MFK, Bordage G, Norman G. Interactive voice response to assess residents' laparoscopic skills: An instrument validation study. In: Annual Meeting of the Council-on-Resident-Education-in-Obstetrics-and-Gynecology/Association-o f-Professors-of-Gynecology-and-Obstetrics; 2003 Mar 01-04; Anaheim, California; 2003. p. 674-678.
- Gumbs AA, Hogle NJ, Fowler DL. Evaluation of resident laparoscopic performance using Global Operative Assessment of Laparoscopic Skills. J Am Coll Surg 2007;204(2): 308-313.
- 34. Sidhu RS, Vikis E, Cheifetz R, Phang T. Self-assessment during a 2-day laparoscopic colectomy course: can surgeons judge how. well they are learning new skills? In: 92nd Annual Meeting of the North-Pacific-Surgical-Association; 2005 Nov 11-12; Vancouver, Canada; 2005. p. 677-680.
- 35. Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondre K, Stanbridge D, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg* 2005;**190**(1): 107-113.
- 36. McCluney AL, Vassiliou MC, Kaneva PA, Cao J, Stanbridge DD, Feldman LS, et al. FLS simulator performance predicts intraoperative laparoscopic skill. *Surg Endosc* 2007;**21**(11): 1991-1995.
- Chang L, Hogle NJ, Moore BB, Graham MJ, Sinanan MN, Bailey R, et al. Reliable assessment of laparoscopic performance in the operating room using videotape analysis. *Surg Innov* 2007;**14**(2): 122-126.
- 38. Ault G, Reznick R, MacRae H, Leadbetter W, DaRosa D, Joehl R, et al. Exporting a technical skills evaluation technology to other sites. *Am J Surg* 2001;**182**(3): 254-256.
- 39. Beard JD, Choksy S, Khan S. Assessment of operative competence during carotid endarterectomy. *Br J Surg* 2007;**94**(6): 726-730.
- Bodle JF, Kaufmann SJ, Bisson D, Nathanson B, Binney DM. Value and face validity of objective structured assessment of technical skills (OSATS) for work based assessment of surgical skills in obstetrics and gynaecology. *Med Teach* 2008;**30**(2): 212-216.

- Faulkner H, Regehr G, Martin J, Reznick R. Validation of an objective structured assessment of technical skill for surgical residents. *Acad Med* 1996;**71**(12): 1363-1365.
- 42. Fialkow M, Mandel L, VanBlaricom A, Chinn M, Lentz G, Goff B. A curriculum for Burch colposuspension and diagnostic cystoscopy evaluated by an objective structured assessment of technical skills. In: Annual Meeting of the Council-on-Resident-Education-in-Obstetrics-and-Gynecology/ Association-o f-Professors-of-Gynecology-and-Obstetrics; 2007 Mar 07-10; Salt Lake City, UT; 2007.
- Goff B, Mandel L, Lentz G, Vanblaricom A, Oelschlager AM, Lee D, et al. Assessment of resident surgical skills: is testing feasible? *Am J Obstet Gynecol* 2005;**192**(4): 1331-1338; discussion 1338-1340.
- 44. Goff BA, Lentz GM, Lee D, Fenner D, Morris J, Mandel LS. Development of a bench station objective structured assessment of technical skills. *Obstet Gynecol* 2001;**98**(3): 412-416.
- Goff BA, Lentz GM, Lee D, Houmard B, Mandel LS. Development of an objective structured assessment of technical skills for obstetric and gynecology residents. *Obstet Gynecol* 2000;**96**(1): 146-150.
- Goff BA, Nielsen PE, Lentz GM, Chow GE, Chalmers RW, Fenner D, et al. Surgical skills assessment: A blinded examination of obstetrics and gynecology residents. *Am J Obstet Gynecol* 2002;**186**(4): 613-617.
- Goff BA, VanBlaricom A, Mandel L, Chinn M, Nielsen P. Comparison of objective, structured assessment of technical skills with a virtual reality hysteroscopy trainer and standard latex hysteroscopy model. In: Annual Meeting of the Association-of-Professors-of-Gynecology-and-Obstetrics/ Council-on-Reside nt-Education-in-Obstetrics-and-Gynecology; 2005 Mar 02-06; Salt Lake City, UT; 2005. p. 407-412.
- 48. Larsen CR, Grantcharov T, Schouenborg L, Ottosen C, Soerensen JL, Ottesen B. Objective assessment of surgical competence in gynaecological laparoscopy: development and validation of a procedure-specific rating scale. *BJOG* 2008;**115**(7): 908-916.
- 49. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997;**84**(2): 273-278.
- 50. Nielsen PE, Foglia LM, Mandel LS, Chow GE. Objective structured assessment of technical skills for episiotomy repair. *Am J Obstet Gynecol* 2003;**189**(5): 1257-1260.
- 51. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. Testing technical skill via an innovative "bench station" examination. *Am J Surg* 1997;**173**(3): 226-230.
- 52. Shah J, Munz Y, Manson J, Moorthy K, Darzi A. Objective assessment of small bowel anastomosis skill in trainee general surgeons and urologists. *World J Surg* 2006;**30**(2): 248-251.
- Siddiqui NY, Stepp KJ, Lasch SJ, Mangel JM, Wu JM. Objective structured assessment of technical skills for repair of fourth-degree perineal lacerations. In: 34th Annual Meeting of the Society-of-Gynecological-Surgeons; 2008 Apr 14-16; Savannah, GA; 2008.
- VanBlaricom AL, Goff BA, Chinn M, Icasiano MM, Nielsen P, Mandel L. A new curriculum for hysteroscopy training as demonstrated by an objective structured assessment of technical skills (OSATS). *Am J Obstet Gynecol* 2005;**193**(5): 1856-1865.
- 55. Winckel CP, Reznick RK, Cohen R, Taylor B. Reliability and construct validity of a structured technical skills assessment form. *Am J Surg* 1994;**167**(4): 423-427.
- 56. 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;**245**(6): 992-999.

- 57. Broe D, Ridgway PF, Johnson S, Tierney S, Conlon KC. Construct validation of a novel hybrid surgical simulator. *Surg Endosc* 2006;**20**(6): 900-904.
- Brydges R, Sidhu R, Park J, Dubrowski A. Construct validity of computer-assisted assessment: quantification of movement processes during a vascular anastomosis on a live porcine model. *Am J Surg* 2007;**193**(4): 523-529.
- 59. Datta V, Chang A, Mackay S, Darzi A. The relationship between motion analysis and surgical technical assessments. *Am J Surg* 2002;**184**(1): 70-73.
- Scott DJ, Rege RV, Bergen PC, Guo WA, Laycock R, Tesfay ST, et al. Measuring operative performance after laparoscopic skills training: edited videotape versus direct observation. *J Laparoen-dosc Adv Surg Tech A* 2000;**10**(4): 183-190.
- 61. Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med* 1998;**73**(9): 993-997.
- 62. Scott DJ, Bergen PC, Rege RV, Laycock R, Tesfay ST, Valentine RJ, et al. Laparoscopic training on bench models: better and more cost effective than operating room experience? *J Am Coll Surg* 2000;**191**(3): 272-283.
- Bann SD, Khan MS, Darzi AW. Measurement of surgical dexterity using motion analysis of simple bench tasks. *World J Surg* 2003;27(4): 390-394.
- 64. Brydges R, Classen R, Larmer J, Xeroulis G, Dubrowski A. Computer-assisted assessment of onehanded knot tying skills performed within various contexts: a construct validity study. *Am J Surg* 2006;**192**(1): 109-113.
- 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. *J Am Coll Surg* 2001;**193**(5): 479-485.
- 66. Egi H, Okajima M, Yoshimitsu M, Ikeda S, Miyata Y, Masugami H, et al. Objective assessment of endoscopic surgical skills by analyzing direction-dependent dexterity using the Hiroshima University Endoscopic Surgical Assessment Device (HUESAD). *Surg Today* 2008;**38**(8): 705-710.
- 67. Francis NK, Hanna GB, Cuschieri A. Reliability of the Advanced Dundee Endoscopic Psychomotor Tester for bimanual tasks. *Arch Surg* 2001;**136**(1): 40-43.
- 68. Francis NK, Hanna GB, Cuschieri A. The performance of master surgeons on the Advanced Dundee Endoscopic Psychomotor Tester: contrast validity study. *Arch Surg* 2002;**137**(7): 841-844.
- Macmillan AI, Cuschieri A. Assessment of innate ability and skills for endoscopic manipulations by the Advanced Dundee Endoscopic Psychomotor Tester: predictive and concurrent validity. *Am J Surg* 1999;**177**(3): 274-277.
- Pellen M, Horgan L, Roger Barton J, Attwood S. Laparoscopic surgical skills assessment: can simulators replace experts? World J Surg 2009;33(3): 440-447.
- Pellen MG, Horgan LF, Barton JR, Attwood SE. Construct validity of the ProMIS laparoscopic simulator. Surg Endosc 2009;23(1): 130-139.
- 72. Van Sickle KR, McClusky DA, 3rd, Gallagher AG, Smith CD. Construct validation of the ProMIS simulator using a novel laparoscopic suturing task. *Surg Endosc* 2005;**19**(9): 1227-1231.
- 73. Smith SG, Torkington J, Brown TJ, Taffinder NJ, Darzi A. Motion analysis. *Surg Endosc* 2002;**16**(4): 640-645.
- 74. Oostema JA, Abdel MP, Gould JC. Time-efficient laparoscopic skills assessment using an augmented-reality simulator. *Surg Endosc* 2008;**22**(12): 2621-2624.
- 75. Datta V, Bann S, Mandalia M, Darzi A. The surgical efficiency score: a feasible, reliable, and valid method of skills assessment. *Am J Surg* 2006;**192**(3): 372-378.

- 76. Chmarra MK, Klein S, de Winter JC, Jansen FW, Dankelman J. Objective classification of residents based on their psychomotor laparoscopic skills. *Surg Endosc* 2009.
- 77. Andreatta PB, Woodrum DT, Gauger PG, Minter RM. LapMentor metrics possess limited construct validity. *Simul Healthc* 2008;**3**(1): 16-25.
- Cope DH, Fenton-Lee D. Assessment of laparoscopic psychomotor skills in interns using the mist virtual reality simulator: A prerequisite for those considering surgical training? *ANZ J Surg* 2008;**78**(4): 291-296.
- 79. 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;**19**(3): 401-405.
- 80. Eriksen JR, Grantcharov T. Objective assessment of laparoscopic skills using a virtual reality stimulator. *Surg Endosc* 2005;**19**(9): 1216-1219.
- Gallagher AG, Richie K, McClure N, McGuigan J. Objective psychomotor skills assessment of experienced, junior, and novice laparoscopists with virtual reality. *World J Surg* 2001;25(11): 1478-1483.
- 82. Gallagher AG, Satava RM. Virtual reality as a metric for the assessment of laparoscopic psychomotor skills Learning curves and reliability measures. *Surg Endosc* 2002;**16**(12): 1746-1752.
- 83. Grantcharov TP, Rosenberg J, Pahle E, Funch-Jensen P. Virtual reality computer simulation. *Surg Endosc* 2001;**15**(3): 242-244.
- 84. Kundhal PS, Grantcharov TP. Psychomotor performance measured in a virtual environment correlates with technical skills in the operating room. *Surg Endosc* 2009;**23**(3): 645-649.
- 85. Langelotz C, Kilian M, Paul C, Schwenk W. LapSim virtual reality laparoscopic simulator reflects clinical experience in German surgeons. *Langenbecks Arch Surg* 2005;**390**(6): 534-537.
- 86. Larsen CR, Grantcharov T, Aggarwal R, Tully A, Sorensen JL, Dalsgaard T, et al. Objective assessment of gynecologic laparoscopic skills using the LapSimGyn virtual reality simulator. *Surg Endosc* 2006;**20**(9): 1460-1466.
- Madan AK, Frantzides CT, Sasso LM. Laparoscopic baseline ability assessment by virtual reality. J Laparoendosc Adv Surg Tech A 2005;15(1): 13-17.
- McDougall EM, Corica FA, Boker JR, Sala LG, Stoliar G, Borin JF, et al. Construct validity testing of a laparoscopic surgical simulator. J Am Coll Surg 2006;202(5): 779-787.
- 89. McNatt SS, Smith CD. A computer-based laparoscopic skills assessment device differentiates experienced from novice laparoscopic surgeons. *Surg Endosc* 2001;**15**(10): 1085-1089.
- Rosenthal R, Gantert WA, Hamel C, Hahnloser D, Metzger J, Kocher T, et al. Assessment of construct validity of a virtual reality laparoscopy simulator. *J Laparoendosc Adv Surg Tech A* 2007;**17**(4): 407-413.
- 91. Rosenthal R, Gantert WA, Scheidegger D, Oertli D. Can skills assessment on a virtual reality trainer predict a surgical trainee's talent in laparoscopic surgery? *Surg Endosc* 2006;**20**(8): 1286-1290.
- 92. Schijven M, Jakimowicz J. Construct validity: experts and novices performing on the Xitact LS500 laparoscopy simulator. *Surg Endosc* 2003;**17**(5): 803-810.
- Sherman V, Feldman LS, Stanbridge D, Kazmi R, Fried GM. Assessing the learning curve for the acquisition of laparoscopic skills on a virtual reality simulator. Surg Endosc 2005;19(5): 678-682.
- 94. Taffinder N, 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.
- 95. 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;**20**(3): 511-518.

- 96. Verdaasdonk EG, Stassen LP, Schijven MP, Dankelman J. Construct validity and assessment of the learning curve for the SIMENDO endoscopic simulator. *Surg Endosc* 2007;**21**(8): 1406-1412.
- 97. Woodrum DT, Andreatta PB, Yellamanchilli RK, Feryus L, Gauger PG, Minter RM. Construct validity of the LapSim laparoscopic surgical simulator. *Am J Surg* 2006;**191**(1): 28-32.
- 98. Yamaguchi S, Konishi K, Yasunaga T, Yoshida D, Kinjo N, Kobayashi K, et al. Construct validity for eye-hand coordination skill on a virtual reality laparoscopic surgical simulator. *Surg Endosc* 2007;**21**(12): 2253-2257.
- 99. Zhang A, Hunerbein M, Dai Y, Schlag PM, Beller S. Construct validity testing of a laparoscopic surgery simulator (Lap Mentor): evaluation of surgical skill with a virtual laparoscopic training simulator. *Surg Endosc* 2008;**22**(6): 1440-1444.
- 100. 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;**197**(5): e1-e4.
- 101. 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;**96**(9): 1086-1093.
- Aggarwal R, Grantcharov TP, Eriksen JR, Blirup D, Kristiansen VB, Funch-Jensen P, et al. An evidencebased virtual reality training program for novice laparoscopic surgeons. *Ann Surg* 2006;**244**(2): 310-314.
- Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. *Am J Surg* 2006;**191**(1): 128-133.
- 104. Beard JD. Setting standards for the assessment of operative competence. In: 18th Annual Meeting of the European-Society-for-Vascular-Surgery; 2004 Sep 17-19; Innsbruck, Austria; 2004. p. 215-218.
- 105. Beard JD, Jolly BC, Newble DI, Thomas WE, Donnelly J, Southgate LJ. Assessing the technical skills of surgical trainees. *Br J Surg* 2005;**92**(6): 778-782.
- Driscoll PJ, Paisley AM, Paterson-Brown S. Video assessment of basic surgical trainees' operative skills. Am J Surg 2008;196(2): 265-272.
- 107. Haddad M, Zelikovski A, Gutman H, Haddad E, Reiss R. Assessment of surgical residents competence based on postoperative complications. *Int Surg* 1987;**72**(4): 230-232.
- Leong JJ, Nicolaou M, Atallah L, Mylonas GP, Darzi AW, Yang GZ. HMM assessment of quality of movement trajectory in laparoscopic surgery. *Comput Aided Surg* 2007;**12**(6): 335-346.
- 109. Rosen J, Hannaford B, Richards CG, Sinanan MN. Markov modeling of minimally invasive surgery based on tool/tissue interaction and force/torque signatures for evaluating surgical skills. *IEEE Trans Biomed Eng* 2001;**48**(5): 579-591.
- 110. Rosen J, MacFarlane M, Richards C, Hannaford B, Sinanan M. Surgeon-tool force/torque signatures evaluation of surgical skills in minimally invasive surgery. *Stud Health Technol Inform* 1999;**62**: 290-296.
- 111. Rosen J, Solazzo M, Hannaford B, Sinanan M. Objective laparoscopic skills assessments of surgical residents using Hidden Markov Models based on haptic information and tool/tissue interactions. *Stud Health Technol Inform* 2001;**81**: 417-423.
- 112. Tang B, Hanna GB, Carter F, Adamson GD, Martindale JP, Cuschieri A. Competence assessment of laparoscopic operative and cognitive skills: Objective Structured Clinical Examination (OSCE) or Observational Clinical Human Reliability Assessment (OCHRA). *World J Surg* 2006;**30**(4): 527-534.

- 113. Van Sickle KR, Baghai M, Huang IP, Goldenberg A, Smith CD, Ritter EM. Construct validity of an objective assessment method for laparoscopic intracorporeal suturing and knot tying. *Am J Surg* 2008;**196**(1): 74-80.
- Wilasrusmee C, Lertsithichai P, Kittur DS. Vascular anastomosis model: relation between competency in a laboratory-based model and surgical competency. *Eur J Vasc Endovasc Surg* 2007;**34**(4): 405-410.
- Derossis AM, Antoniuk M, Fried GM. Evaluation of laparoscopic skills: A 2-year follow-up during residency training. *Can J Surg* 1999;**42**(4): 293-296.
- 116. Derossis AM, Fried GM, Abrahamowicz M, Sigman HH, Barkun JS, Meakins JL. Development of a model for training and evaluation of laparoscopic skills. *Am J Surg* 1998;**175**(6): 482-487.
- 117. Feldman LS, Hagarty SE, Ghitulescu G, Stanbridge D, Fried GM. Relationship between objective assessment of technical skills and subjective in-training evaluations in surgical residents. *J Am Coll Surg* 2004;**198**(1): 105-110.
- Fraser SA, Klassen DR, Feldman LS, Ghitulescu GA, Stanbridge D, Fried GM. Evaluating laparoscopic skills - Setting the pass/fail score for the MISTELS system. *Surg Endosc* 2003;**17**(6): 964-967.
- 119. Fried GM, Derossis AM, Bothwell J, Sigman HH. Comparison of laparoscopic performance in vivo with performance measured in a laparoscopic simulator. *Surg Endosc* 1999;**13**(11): 1077-1081; discussion 1082.
- 120. 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;**240**(3): 518-525; discussion 525-518.
- 121. Swanstrom LL, Fried GM, Hoffman KI, Soper NJ. Beta test results of a new system assessing competence in laparoscopic surgery. *J Am Coll Surg* 2006;**202**(1): 62-69.
- 122. Vassiliou MC, Ghitulescu GA, Feldman LS, Stanbridge D, Leffondre K, Sigman HH, et al. The MIS-TELS program to measure technical skill in laparoscopic surgery : evidence for reliability. *Surg Endosc* 2006;**20**(5): 744-747.

CHAPTER 2.2

An intermediate curriculum for advanced laparoscopic skills training with virtual reality simulation

H.W.R. Schreuder¹, P.D. van Hove^{2,3}, J.A. Janse¹, R.H.M. Verheijen¹, L.P.S. Stassen⁴, J. Dankelman²

 Department of Gynaecologic Surgery and Oncology, University Medical Centre Utrecht, the Netherlands
 Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands
 Department of Surgery, Reinier de Graaf Group, Delft, the Netherlands
 Department of Surgery, Maastricht University Medical Center, Maastricht, the Netherlands

Published in: J Minim Invasive Gynecol. 2011 Sep-Oct; 18(5):597-606

ABSTRACT

Study Objective To estimate face- and construct validity for a novel curriculum designed for intermediately skilled laparoscopic surgeons on the Simendo[®] virtual reality simulator. It consists of five exercises, focusing on training precision and coordination between both hands.

Setting Three University hospitals and 4 teaching hospitals in the Netherlands

Subjects Residents, consultants and laparoscopic experts (N=69) in the fields of general surgery, gynecology and urology participated. Participants were divided into four groups based on their level of laparoscopic experience: 'residents year 1-3' (N=15), 'residents year 4-6' (N=17), 'consultants' (N=19) and 'laparoscopic experts' (N=18).

Interventions Participants completed three runs of five exercises. The first run was an introduction and the second and third run were used for analysis. The parameters 'time', 'path length', 'collisions' and 'displacement' were compared between groups. Afterwards the participants completed a questionnaire to evaluate their laparoscopic experience and identify issues concerning the simulator and exercises.

Results The expert group was significantly faster (p<.05) than other groups in 4 of 5 exercises. The parameter displacement demonstrated a significant difference between the expert group and other groups in two of the four exercises in which this parameter was relevant (p<0.05). In the questionnaire (N=68), training capacity of the curriculum was scored with a median of 4 points on a 5-point Likert scale. Of all participants, 92.6% indicated that this curriculum is suitable as an addition to a basic skills module within their residency program.

Conclusion Face- and construct validity were estimated for an advanced virtual reality curriculum for intermediately skilled laparoscopic surgeons. The results indicate that the curriculum is suitable for training of residents and consultants and to assess and maintain their laparoscopic skills.

INTRODUCTION

Since its introduction, laparoscopic surgery has become of great importance to the practice of surgical care. However, laparoscopic surgery requires different skills compared to open procedures, leading to a recognized learning curve.¹ Laparoscopic surgery requires distinct psychomotor abilities, ambidextrous and hand-eye coordination, depth perception, and manipulation of delicate structures with limited haptic sensation.

It is now widely accepted that skills training outside the operating room is essential for residents. The use of animal models is prohibited in several countries. Box trainers and video trainers have the disadvantage of the lack of automated performance assessment. In comparison, virtual reality (VR) trainers provide a safe and standardized environment to practice specific skills and simultaneously measure objectively the performance of the trainee.² VR training can supplement standard laparoscopic box (video) training and is at least as effective.^{3,4} In recent years, several VR trainers have been validated for training in general surgery^{5,6}, urology⁷ and gynecology^{8,9} and a significant correlation between operative performance and psychomotor performance on VR simulators has been demonstrated.¹⁰ The acquired skills on a VR simulator are not procedure specific. Rather, VR training improves overall laparoscopic surgical skills.¹¹ Skills acquired on a VR simulator are transferable to actual laparoscopic operations in animals and in human patients.¹²⁻¹⁷ Implementation of standardized VR training curricula in residency training programs is preferred in order to acquire a predetermined level of proficiency before progression to the operating theatre.^{15,18,19} However, most VR curricula focus on basic laparoscopic skills and are suited for junior residents, whereas senior residents prefer live animal training and training on box/video trainers for more advanced skills training.^{20,21} With attention being raised for continuous training and assessment²⁰, a more complex training program is needed. These more complex training programs should focus on ambidextrous skill development and precision of instrument handling, aiming at the more experienced residents and surgeons. For this study, a set of VR exercises was developed to train more advanced laparoscopic skills. The exercises provide a new challenge when the trainee has succeeded the basic exercises. This provides increasing levels of training difficulty and practice variety, which will improve the retention and transfer of simulator-acquired skills.²¹ The aim of the present study was to evaluate faceand construct validity for the developed set of exercises for training and assessment of advanced laparoscopic skills on this VR simulator.

MATERIALS AND METHODS

Curriculum development

The new exercises for the SIMENDO[°] Virtual Reality simulator (Simendo, Rotterdam, the Netherlands) were developed during the six months prior to the study. In the development phase a gynecologist and a general surgeon worked closely together with the computer programmer to design exercises appropriate for surgeons and residents who have passed the initial learning curve of laparoscopic surgery and already perform laparoscopic procedures without assistance. This resulted in an "intermediate curriculum" consisting of five exercises (Fig. 1, number 2 to 6). Every exercise was designed to train precision, stability, ambidextrous coordination and co-operation between the left and the right instrument. The exercises, their training goals and the measured parameters are described in Table 1.



Figure 1 Handles of the Virtual Reality Trainer and the 5 exercises of the curriculum: 'Slide & Drop' (2), 'Clip a vessel' (3), 'Ring and Needle' (4), 'The Carrousel' (5) and 'Single Knot' (6).

Participants

Surgeons and residents in gynecology, urology and general surgery from three regions in the Netherlands were recruited for voluntary participation (n=69). Four groups were formed based on laparoscopic experience and status. Group 1 (n=15) consisted of post-graduate residents year 1-3 (PGY 1-3), group 2 (n=17) of post-graduate residents year 4-6 (PGY 4-6), group 3 (n=19) consisted of consultants and group 4 (n=18) of expert laparoscopic surgeons. Consultants and laparoscopic experts were differentiated based on the number of selected "advanced laparoscopic" procedures performed. For gynecologists the selected procedures were: laparoscopic hysterectomy, laparoscopic sacrocolpopexie and laparoscopic lymphadenectomy. For general surgeons the selected procedures were: laparoscopic colectomy and laparoscopic colectomy and laparoscopic bariatric procedures. For urologists laparoscopic prostatectomy was selected as

| Exercise no. name | Exercise description | Training goals | Parameter |
|-----------------------|--|---|---|
| 1. Slide and drop | Dropping of 4 balls into holes under slides, which need to be moved by the opposite hand | Ambidextrous coordination | Task time, path length, displacement (change in location of the floor and slides) |
| 2. Clip a vessel | Applying 4 clips at designated places while completely grasping the vessel | Ambidextrous coordination and precision training | Task time, path length, precision of clip application, displacement (change in location of the vessel) |
| 3. Ring and Needle | Putting a needle through a ring thrice, while keeping the ring stable | Ambidextrous coordination, stability and precision training | Task time, path length, displacement (change in location of the ring) |
| 4. The Carrousel | Grasping 3 dices of a carousel and stacking them as a tower on a small plateau | Ambidextrous coordination, stability and precision training | Task time, path length, stacked dices |
| 5. Single Knot | Tying a single knot around a bar, while keeping the bar stable | Ambidextrous coordination and stability | Task time, path length and displacement (change in location of the bar), collisions (too much pressure on the floor' |

 Table 1
 Overview of the five exercises in the 'intermediate curriculum'.

an advanced procedure. To be considered an expert the number of performed advanced procedures should be more than 10 for at least two of the selected procedures, or more than 50 for one of the selected procedures to be considered an expert. Characteristics of the different groups are shown in Table 2. All participants where asked for their prior experience with laparoscopic skills training. Experience with box-trainers (video-trainers), virtual reality trainers and live-animal training was estimated in hours.

Equipment

The SIMENDO[°] Virtual Reality simulator (Simendo, Rotterdam, The Netherlands) was used. This system consists of a software interface and two hardware laparoscopic instruments (Fig. 1, number 1) connected with a USB plug to a laptop computer (Acer[°] Aspire 5924G). The laptop contains an Intel[®] Core[™] 2 Duo processor T5550 (1.83 GHz, 667 MHz FSB, 2 MB L2 cache), TurboCache[™] 3 GB DDR2, with graphical card NVIDIA[°] GeForce[™] 8600 GT, 15.4″ WXGA Acer Crystalbyte[™] LCD monitor and software Microsoft Windows XP[™].

Face validity

Face validity is defined as the extent to which the simulation (or test) resembles the experience in the real world.²² To investigate this, participants filled out a questionnaire immediately after performing the five exercises on the simulator. The participant's demographics, laparoscopic training experience and laparoscopic theatre experience

| Total n = 69 | Group 1 residents PGY 1-3 (n=15) | Group 2 residents PGY 4-6 (n=17) | Group 3 consultants (n=19) | Group 4 laparoscopic experts (n=18) |
|------------------------------------|--|--|--|---|
| Gender Male/ | M: n=6 | M: n=12 | M: n=12 | M: n=15 |
| Female | F: n=9 | F: n=5 | F: n=7 | F: n=3 |
| Mean age (range) | 30 (26-35) | 33 (29-37) | 46 (35-63) | 45 (33-58) |
| Criteria | Resident PGY 1-3 | Resident PGY 4-6 | non or < 50 advanced procedures* | > 50 advanced procedures* |
| Distribution within | 1st year: n=1 | 4th year: n=6 | years consultant | years consultant |
| group | 2nd year: n=6 | 5th year: n=6 | ≤5 year: n=7 | ≤5 year: n=3 |
| | 3rd year: n=8 | 6th year: n=5 | >5 year: n=12 | >5year: n=12 |
| | | | >10 year: n=7 | >10year: n=5 |
| | | | >20 year: n=4 | >20 year: n=2 |
| Specialism | Sur: n=7 | Sur: n=6 | Sur: n=6 | Sur: n=11 |
| | Gyn: n=8 | Gyn: n=5 | Gyn: n=11 | Gyn: n=6 |
| | Uro: n=0 | Uro: n=6 | Uro: n=2 | Uro: n=1 |
| Dominant hand | R: n=13 | R: n=15 | R: n=18 | R: n=18 |
| | L: n=2 | L: n=2 | L: n=1 | L: n=0 |
| Can tie a knot Iaparoscopically | n=6 | n=12 | n=13 | n=18 |

Table 2 Characteristics of the different groups.

*advanced procedures in laparoscopic gynecology are: hysterectomy, sacrocolpopexie, pelvic lymfadenectomy - in surgery: Nissen fundoplication, bariatric procedures, colectomy - in urology: prostatectomy.

were evaluated. Additionally, the opinion of each participant about the simulator was asked. The first section of the questionnaire contained five questions on the realism of the simulator; the second section consisted of ten questions regarding the training capacities in general and the suitability for training residents or surgeons. In the third section, six questions were asked to evaluate each exercise separately. The questions were presented on a 5-point Likert scale.²³ Finally, two statements were given concerning the addition of the "intermediate curriculum" to the existing "novice curriculum" of this simulator and the implementation of this simulator in the current residency training programs. These could be answered with "agree", "disagree" or "no opinion".

Construct validity

Construct validity is defined as the extent to which a simulation (test) is able to identify the quality, ability or trait it is designed to measure. This simulation addresses technical skills and it is therefore expected that it differentiates between a good (expert) and a bad performer (novice).²² To investigate construct validity the participants performed three consecutive repetitions of each of the five exercises of the curriculum. The exercises were 'Slide and Drop' 'Clip a Vessel', 'Ring and Needle', 'The Carrousel' and 'Single Knot'. Before starting on the simulator, the exercises were briefly explained by the test supervisor,

and introduced with an instruction video for each exercise. The first run of each exercise was used to familiarize the participant with the simulator, and verbal instructions were given whenever necessary. The second and third repetitions were used for analysis and participants received minimal or no guidance during these runs. To evaluate construct validity the following parameters were compared between the different groups for each exercise separately: 1) task time, 2) path length, 3) displacement and 4) collisions.

Use of statistics

Data were analyzed using the statistical software package SPSS 16.0 (SPSS Inc, Chicago, IL). Differences in performance between groups were analyzed using the Kruskal-Wallis test. If the Kruskal-Wallis Test resulted in a significant difference, then a separate analysis was performed, comparing the expert group with the other groups, using the Mann-Whitney test with post-hoc Dunn's (Bonferroni) correction. To verify the minimum sample size a power analysis was performed. A total sample of 68 subjects achieves a power of 0.81 using the Kruskal-Wallis Test with a target significance level of 0.05. The average within group standard deviation assuming the alternative distribution is 1.0. (PASS 2008 NCSS, LCC. Kayville, UT). A level of p<.05 was considered to be statistically significant. Values are presented as medians unless stated otherwise.

RESULTS

Prior Training

Nearly all residents and experts, but only 50% of the consultants, had experience on box-trainers. Most experts had completed significantly more training hours on box-trainers than the other groups. Approximately half of the participants had experience on VR trainers. Several participants had prior experience with the virtual reality simulator used in this study. These participants were equally divided among the four groups, and therefore had no significant influence on the statistical group comparison as a whole. The majority of the residents did not have experience with live animal training in contrast to the experts, who nearly all had a history of life animal training (Fig. 2).

Face validity

Of the 69 participants 68 participants (97%) filled out the questionnaire. Table 3 summarizes the median values of the scores considering the realism and training capacity of the curriculum. The training capacity of the curriculum in general was appreciated with a median score of 4.0. Hand-eye coordination and the training capacity of cooperation of the right and left hand were both rated with a 4.0. The suitability of the intermediate curriculum to train PGY 1-3, PGY 4-6, consultants and expert laparoscopic surgeons



Figure 2 Training experience of the participants (n=69) on box-trainers, the SIMENDO virtual reality trainer, other virtual reality trainers and live animals.

was rated 5.0, 4.0, 4.0 and 3.0, respectively. The results for face validity by exercise are shown in Table 4. The realism of depth perception received the lowest score (3.0). The lack of haptic feedback was scored as 'disturbing' in two exercises, receiving relatively low scores. Despite these drawbacks, 92.6% of the participants agreed that the new curriculum for this virtual reality trainer was suitable for implementation in the current resident training programs. The remaining 7.4% of the participants answered this question with 'no opinion'.

Construct validity

All of the 69 participants completed the three repetitions of the five exercises. Median values of the measured parameters in the second and third run for all exercises are shown in Table 5. The parameter 'task time' was significantly different between the groups in four exercises. Experts (group 4) performed significantly faster compared to one or more

| Question | Group 1 residents PGY 1-3 (n=15) | Group 2 residents PGY 4-6 (n=17) | Group 3 consultants (n=19) | Group 4 laparoscopic experts (n=17) | Overall (n=68) |
|---|---|---|----------------------------------|--|-------------------|
| What do you think of the realism of the curriculum concerning? (not realisticvery realistic) | | | | | |
| the appearance of the instruments | 4.0 (4.0 – 5.0) | 4.0 (4.0 – 5.0) | 4.0 (3.0 – 5.0) | 4.0 (4.0 – 5.0) | 4.0 (4.0 – 5.0) |
| the movements of the instruments | 4.0 (4.0 – 4.0) | 4.0 (3.5 – 4.0) | 3.5 (3.0 – 4.0) | 3.0 (3.0 – 4.0) | 4.0 (3.0 – 4.0) |
| freedom of movements of the instruments | 4.0 (3.0 – 4.0) | 4.0 (3.5 – 4.0) | 4.0 (3.0 – 4.0) | 4.0 (3.0 – 4.0) | 4.0 (3.0 – 4.0) |
| depth perception | 3.0 (2.0 – 3.0) | 2.0 (2.0 – 3.0) | 2.0 (2.0 – 3.0) | 3.0 (1.5 – 3.0) | 3.0 (2.0 – 3.0) |
| interaction of the instruments with other objects | 3.0 (2.0 – 4.0) | 3.0 (2.0 – 4.0) | 3.0 (2.0 – 3.0) | 3.0 (2.0 – 3.0) | 3.0 (2.0 – 3.0) |
| What do you think of the training capacity of the curriculum? (very badvery good) | | | | | |
| in general | 4.0 (4.0 - 5.0) | 4.0 (4.0 – 4.0) | 4.0 (4.0 - 5.0) | 4.0 (3.0 - 4.0) | 4.0 (4.0 – 4.0) |
| eye-hand coordination | 4.0 (4.0 - 5.0) | 4.0 (4.0 – 5.0) | 5.0 (4.0 - 5.0) | 5.0 (4.0 - 5.0) | 4.0 (4.0 – 5.0) |
| depth perception | 3.0 (2.0 – 4.0) | 2.0 (2.0 – 3.0) | 3.0 (2.0 – 4.0) | 2.0 (1.5 – 3.0) | 3.0 (2.0 – 3.0) |
| instrument navigation in general | 4.0 (4.0 - 4.0) | 4.0 (3.0 – 4.0) | 4.0 (3.0 – 5.0) | 4.0 (3.0 – 5.0) | 4.0 (3.0 – 4.0) |
| training left and right hand separately | 4.0 (4.0 – 5.0) | 4.0 (4.0 – 4.5) | 4.0 (4.0 – 5.0) | 4.0 (3.0 – 5.0) | 4.0 (4.0 – 5.0) |
| training co-operation between left and right hand (multiple tasking) | 4.0 (4.0 – 5.0) | 4.0 (4.0 – 5.0) | 4.0 (3.0 – 5.0) | 4.0 (4.0 – 5.0) | 4.0 (4.0 – 5.0) |
| The intermediate curriculum is suitable to train (not suitablevery suitable) | | | | | |
| Residents PGY 1 to 3 | 5.0 (4.0 - 5.0) | 4.0 (4.0 – 5.0) | 5.0 (4.0 - 5.0) | 4.0 (3.0 – 5.0) | 5.0 (4.0 – 5.0) |
| Residents PGY 4 to 6 | 4.0 (4.0 – 4.0) | 4.0 (4.0 – 5.0) | 5.0 (4.0 - 5.0) | 4.0 (3.5 – 5.0) | 4.0 (4.0 – 5.0) |
| Consultants | 4.0 (4.0 – 4.0) | 4.0 (4.0 – 5.0) | 4.0 (3.0 – 5.0) | 4.0 (2.5 – 4.0) | 4.0 (3.3 – 5.0) |
| Laparoscopic experts | 3.0 (2.0 - 3.0) | 2.0 (2.0 – 3.0) | 3.0 (2.0 – 5.0) | 3.0 (2.0 - 4.0) | 3.0 (2.0 – 4.0) |

Table 3Face validity of the intermediate curriculum as a whole (n=68, median score (interquartilerange) on a 1 – 5 Likert scale).

of the other groups in four of the five exercises (Fig. 3): 'Slide & Drop': group 4 versus group 1, 2 and 3 (p=.009; .045; .009), 'Ring & Needle: group 4 versus group 1 (p=0.027), 'The Carrousel': group 4 versus group 1, 2 and 3 (p=.003; .033; <.001) and 'Single Knot': group 4 versus group 2 (p=.033). In the exercise 'Clip a Vessel' there was no significant difference in task time between the expert group and other groups. Additionally for the

| Question | Slide & Drop | Clip a Vessel | Ring & Needle | The Carrousel | Single Knot |
|---|-----------------|-----------------|------------------|------------------|-----------------|
| Do you think the training goal is reached? (not at allyes for sure) | 4.0 (4.0 – 5.0) | 4.0 (3.0 – 4.0) | 4.0 (3.3 – 4.0) | 4.0 (4.0 – 4.0) | 4.0 (4.0 – 5.0) |
| What do you think of? (1= very bad5= very good) | | | | | |
| the set-up of the exercise | 4.0 (4.0 – 5.0) | 4.0 (4.0 – 4.8) | 4.0 (3.0 – 5.0) | 4.0 (4.0 – 5.0) | 4.0 (4.0 – 5.0) |
| the movements of the instruments | 4.0 (4.0 – 4.0) | 4.0 (3.0 – 5.0) | 4.0 (3.0 – 4.0) | 4.0 (4.0 – 4.0) | 4.0 (4.0 – 5.0) |
| the depth perception | 3.0 (2.0 – 4.0) | 3.0 (3.0 – 4.0) | 3.0 (2.0 – 3.8) | 3.0 (2.0 – 4.0) | 3.0 (2.0 – 4.0) |
| the training capacity of the exercise | 4.0 (4.0 – 4.0) | 4.0 (4.0 – 5.0) | 4.0 (3.0 – 4.0) | 4.0 (4.0 – 4.8) | 4.0 (3.3 – 5.0) |
| The lack of haptic feedback is (very disturbingnot disturbing at all) | 2.0 (2.0 – 3.0) | 2.0 (2.0 – 3.0) | 3.0 (2.0 – 4.0) | 3.0 (2.0 – 4.0) | 3.0 (2.3 – 4.0) |

Table 4Face validity of the individual exercises of the curriculum (n=68, median score (interquartilerange) on a 1 to 5 Likert scale).

'Single Knot' exercise, task time was compared between participants who were already able to tie an intracorporeal knot and those who were not. The latter ones took almost twice as long (66.67 versus 102.13 seconds, p=<.001).

The displacement parameter was significantly different between the expert group and other groups in two of four exercises in which this parameter was relevant (it was not relevant in 'The Carrousel'). The experts (group 4) showed less displacement of the slides in 'Slide & Drop' compared to the consultants (group 3) (p=0.027). In the exercise 'Single Knot' the experts (group 4) showed less displacement of the bar while tying a knot than both resident groups: group 4 versus group 1 and 2 (p=.012; .024).

In the exercise 'The Carroussel' significant differences in path length were observed. The path length of the expert group was significantly shorter compared to the consultants: group 4 versus group 3 (p=0.03). In the other four exercises there were no significant differences in path length. The parameter precision, used in the exercise 'Clip a Vessel' showed no significant difference between the expert group and the other groups. For the parameter collision in the exercise 'Single Knot', the experts (group 4) had significantly less collisions with the floor: group 4 versus group 1 and 2 (p=.009; <.001).

| Table 5 Construct | validity. Median values (interqu | artile range) for all relevant para | meters. | | |
|-------------------|----------------------------------|-------------------------------------|--------------------------|-------------------------|----------|
| | Group 1 | Group 2 n=17 | Group 3 n=10 | Group 4 | p-value* |
| | residents PGY 1-3 | residents PGY 4-6 | consultants | laproscopic experts | |
| Slide & Drop | | | | | |
| Task time | 124.85 (103.59 – 144.02) | 121.74 (105.77 – 154.93) | 135.82 (109.06 – 151.72) | 93.99 (78.32 – 119.61) | .007 |
| Path length (R) | 61.50 (45.00 – 74.50) | 54.50 (45.75 – 64.25) | 71.00 (48.00 – 92.00) | 57.25 (41.13 – 71.00) | .463 |
| Path length (L) | 61.00 (55.00 – 72.00) | 54.00 (48.75 – 68.50) | 74.00 (55.00 – 96.00) | 53.75 (47.13 – 71.88) | .091 |
| Displacement | 5.68 (4.42 – 11.81) | 9.97 (6.13 - 12.18) | 16.18 (10.23 – 20.23) | 8.38 (7.37 – 9.70) | .003 |
| Clip a Vessel | | | | | |
| Task time | 49.40 (32.41 – 67.79) | 39.83 (33.92 – 63.84) | 49.34 (34.59 – 59.68) | 37.37 (28.03 – 53.04) | .406 |
| Path length (R) | 21.00 (11.50 – 24.00) | 15.75 (14.50 – 22.00) | 22.00 (16.25 – 25.25) | 19.75 (15.25 – 25.63) | .313 |
| Path length (L) | 11.50 (10.50 – 14.00) | 14.75 (11.38 – 19.00) | 17.00 (13.50 – 22.50) | 13.75 (10.63 – 19.13) | .110 |
| Displacement | 0.84 (0.78 – 0.88) | 0.84 (0.78 – 0.89) | 0.90 (0.85 – 0.94) | 0.88 (0.82 – 0.92) | .043 |
| Precision** | 0.50 (0.00 – 1.50) | 0.50 (0.13 – 0.50) | 0.50 (0.50 – 1.00) | 0.50 (0.00 – 1.00) | .715 |
| Ring & Needle | | | | | |
| Task time | 173.43 (113.70 – 196.61) | 116.51 (102.77 – 144.41) | 109.69 (95.00 – 173.59) | 108.40 (89.06 – 142.94) | .043 |
| Path length (R) | 91.00 (58.00 – 117.50) | 72.50 (65.25 – 83.75) | 83.50 (69.25 – 96.75) | 88.25 (57.38 – 109.25) | .619 |
| Path length (L) | 83.50 (49.00 – 105.00) | 67.00 (52.00 – 75.50) | 77.50 (59.25 – 87.50) | 72.25 (53.25 – 82.88) | .400 |
| Displacement | 39.74 (25.62 – 58.45) | 35.42 (27.32 – 38.32) | 38.43 (26.76 – 51.17) | 28.70 (25.32 – 31.78) | .160 |
| Carrousel | | | | | |
| Task time | 138.60 (107.93 – 161.22) | 122.11 (100.45 – 149.78) | 155.69 (104.96 – 204.08) | 99.17 (72.84 – 111.89) | .001 |
| Path length (R) | 53.50 (42.50 – 60.13) | 48.50 (44.25 – 60.25) | 73.50 (59.00 – 95.88) | 50.00 (46.75 – 63.00) | <.001 |
| Path length (L) | 37.25 (24.88 – 44.88) | 33.50 (26.50 – 50.25) | 53.00 (27.75 – 58.63) | 31.00 (27.75 – 33.75) | .072 |
| Displacement | | I | | I | ı |

Advanced laparoscopic skills training with virtual reality

| | Group 1 | Group 2 | Group 3 | Group 4 | p-value* |
|-----------------|------------------------------|------------------------------|-----------------------|--------------------------------|----------|
| | n=15 residents PGY 1-3 | n=17 residents PGY 4-6 | n=19 consultants | n=18 laproscopic experts | |
| Single Knot | | | | | |
| Task time | 84.45 (65.93 – 111.78) | 77.01 (69.06 – 97.56) | 60.49 (51.34 – 88.46) | 66.07 (47.43 – 84.47) | .021 |
| Path length (R) | 58.50 (42.00 – 83.50) | 58.50 (52.25 – 84.25) | 47.25 (34.75 – 70.63) | 68.00 (54.25 – 77.75) | .271 |
| Path length (L) | 59.00 (55.50 – 74.00) | 63.00 (46.50 – 87.25) | 61.50 (51.50 – 89.50) | 55.00 (49.50 - 77) | .873 |
| Displacement | 9.03 (7.46 – 11.45) | 9.42 (6.09 – 10.33) | 7.61 (5.96 – 10.71) | 5.33 (4.88 – 7.69) | .015 |
| Collision*** | 0.00 (0.00 – 0.00) | 0.50 (0.00 – 1.00) | 0.50 (0.13 – 0.88) | 0.00 (0.00 – 1.00) | .007 |

 Table 5
 Construct validity. Median values (interquartile range) for all relevant parameters. (Continued)

Task time in seconds, Path length and displacement in arbitrary units, R=right instrument, L=left instrument, *=p-value tor Kruskal-Wallis, non-parametric test, " Number of incorrect applied clips, ***=Number of collisions with the floor



Figure 3 Total task time for the four different groups performing the five exercises (bars are medians, boxes show inter quartile range, whiskers show range, dots are outliers and large horizontal bars indicate statistically significant differences, specified with p-values).

DISCUSSION

In this study, face and construct validity were estimated for a new set of exercises for training and assessment of intermediate laparoscopic skills on a commercially available virtual reality simulator. The vast majority of participants indicated that the virtual reality simulator is a useful tool for training laparoscopic skills to residents, urologists, gynecologists and surgeons. Construct validity was demonstrated by the fact that the intermediate curriculum was able to differentiate between subjects with varying laparoscopic experience. More precisely, the expert laparoscopic surgeons performed significantly faster in exercises which required most intensive cooperation of the right and left hand in comparison to the other groups. This suggests training capacity of the curriculum for these groups of trainees, i.e. not only for laparoscopic novices.

Additionally, for two exercises the experts showed less displacement of structures than other groups. In the 'Slide & Drop' exercise they outperformed the consultant group and in the 'Single Knot' exercise they outperformed both resident groups. This strokes with the thought that expert surgeons are highly efficient in their movements and that this is a skill which distinguishes them from other groups and which could be trained.

Although differences were found, the discriminative properties of the exercises vary. The most discriminative exercises are 'Slide and Drop' and 'The Carrousel'. This can be explained by the fact that in these exercises the left and right instruments need to work closely together for a longer period of time. This coordination skill is more prominent in more advanced laparoscopic procedures and therefore correlates with more experience. The least discriminative exercise is 'Clip a Vessel'. This exercise may not be discriminative because it is a relatively short and easy to perform. Task time was the most discriminative parameter. However, task time has often been subject of discussion when addressing distinction in the level of laparoscopic skills. Task time alone does not provide enough information about accuracy and precision. Therefore, on its own, this parameter should not be considered to distinguish different levels of experience. This means that the construct validity of this curriculum cannot be considered very solid at this stage. Further modification of the exercises could improve their differentiating capability.

Furthermore, it can be argued that some participants would perform better than others due to previous experience on this specific simulator. In one exercise ('Clip a Vessel'), indeed participants with experience on this specific simulator performed better than participants who had no experience on this simulator. However, the distribution of participants with experience with this particular simulator is equal among all groups, so it is not expected that the differences between groups are affected by this previous experience.

Depth perception and the lack of haptic feedback were indicated as limitations of the VR simulator used in this study. In order to enhance depth perception, features such

as shadows were added during development of the software. To clarify the position of structures the shadow is mimicked on the floor of the simulation environment. The lack of haptic feedback or unrealistic haptic feedback in VR simulators is often stated as a major disadvantage of VR simulators in comparison with box-trainers. In this study, most participants confirmed the lack of haptic feedback to be disturbing. A subgroup analysis was done to see if there was any correlation between the answer of a participant to this specific question and his or her results on the simulator. No correlation was found.

Madan et al found no difference in laparoscopic skills acquisition when incorporating VR trainers in a curriculum based on box-trainers.²⁴ There was also no difference in skills improvement when comparing VR trainers and computer-enhanced box-trainers.²⁵ Therefore one could argue whether haptic feedback is really necessary for training basic laparoscopic skills in VR trainers. Haptic feedback in VR simulators might play a more important role in more complex procedural exercises where there is a need to apply force on tissue structures (laparoscopic hysterectomy, laparoscopic myomectomy, laparoscopic cholecystectomy). Chmarra et al showed that VR trainers without haptic feedback should only train tasks where force application is not required.²⁶

In this study, participants emphasized the importance of introduction of VR in a residency laparoscopic training program, confirming findings by other authors.²⁷ However, previous studies have shown that without any form of obligation or assessment, trainees are typically not sufficiently motivated to use VR simulators.²⁸⁻³⁰ To make sure a VR simulator will be used, the exercises need to be implemented in a complete laparoscopic training curriculum and failing of these exercises must have certain consequences. For instance, trainees should not be allowed to perform laparoscopic surgery on patients, at the position of primary surgeon, as long as they fail to reach the pre-set performance levels during training. This will have a positive effect on operating times and patient safety.¹⁵ When incorporating a VR trainer in a laparoscopic training curriculum one should carefully choose the exercises and only use validated exercises, construct validity being the minimum requirement. A complete VR curriculum should have exercises targeting different levels of experience and ideally should contain basic skills, advanced skills and procedural tasks (like salpingectomy, cholecystectomy or appendectomy). Such a curriculum is motivating for participants with different levels of experience. The present study adds an intermediate module to laparoscopic virtual reality training and demonstrates the appreciation by the possible trainees. Adherence to a training curriculum is also influenced by the ease of participation and receiving feedback.

Face- and construct validity have been estimated for the 'intermediate curriculum' on this simulator. Although results for face validity were unequivocally positive, construct validity cannot be considered very solid at this stage. This curriculum might eventually not only be suitable for residents throughout their surgical training, but also for consultants. It would therefore offer a training continuum for residents in addition to basic VR skills training and might also open this training environment for the registered surgeon. To improve construct validity, modification of exercises might be needed and further research should focus on the predictive validity of this simulator to show whether it can predict eventual better performance in the operating room.

REFERENCES

- 1. Figert PL, Park AE, Witzke DB, Schwartz RW. Transfer of training in acquiring laparoscopic skills. *J Am Coll Surg.* 2001;193:533-537.
- 2. Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. *Br J Surg.* 2004;91:1549-1558.
- 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;95:1088-1097.
- 4. Gurusamy KS, Aggarwal R, Palanivelu L, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. *Cochrane Database Syst Rev.* 2009;CD006575.
- 5. Schijven M, Jakimowicz J. Construct validity: experts and novices performing on the Xitact LS500 laparoscopy simulator. *Surg Endosc*. 2003;17:803-810.
- 6. 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;20:511-518.
- Watterson JD, Beiko DT, Kuan JK, Denstedt JD. Randomized prospective blinded study validating acquisition of ureteroscopy skills using computer based virtual reality endourological simulator. *J Urol.* 2002;168:1928-1932.
- Larsen CR, Grantcharov T, Aggarwal R, Tully A, Sorensen JL, Dalsgaard T, Ottesen B. Objective assessment of gynecologic laparoscopic skills using the LapSimGyn virtual reality simulator. Surg Endosc. 2006;20:1460-1466.
- 9. Schreuder HW, van Dongen KW, Roeleveld SJ, Schijven MP, Broeders IA. Face and construct validity of virtual reality simulation of laparoscopic gynecologic surgery. *Am J Obstet Gynecol.* 2009;200:540-548.
- 10. Kundhal PS, Grantcharov TP. Psychomotor performance measured in a virtual environment correlates with technical skills in the operating room. *Surg Endosc.* 2009;23:645-649.
- 11. Lucas SM, Zeltser IS, Bensalah K, Tuncel A, Jenkins A, Pearle MS, Cadeddu JA. Training on a virtual reality laparoscopic simulator improves performance of an unfamiliar live laparoscopic procedure. *J Urol.* 2008;180:2588-2591.
- 12. 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;246:771-779.
- Ahlberg G, Enochsson L, Gallagher AG, Hedman L, Hogman C, McClusky DA, Ramel S, Smith CD, Arvidsson D. Proficiency-based virtual reality training significantly reduces the error rate for residents during first 10 laparoscopic cholecystectomies. *Am J Surg.* 2007;193:797-804.
- 14. Hyltander A, Liljegren E, Rhodin PH, Lonroth H. The transfer of basic skills learned in a laparoscopic simulator to the operating room. *Surg Endosc.* 2002;16:1324-1328.
- Larsen CR, Soerensen JL, Grantcharov TP, Dalsgaard T, Schouenborg L, Ottosen C, Schroeder TV, Ottesen BS. Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. *BMJ*. 2009;338:b1802.
- 16. Schijven MP, Jakimowicz JJ, Broeders IA, Tseng LN. The Eindhoven laparoscopic cholecystectomy training course—improving operating room performance using virtual reality training: results from the first E.A.E.S. accredited virtual reality trainings curriculum. *Surg Endosc*. 2005;19:1220-1226.
- 17. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, Satava RM. Virtual reality training improves operating room performance: results of randomized, double-blinded study. *Ann Surg.* 2002;236:458-463.

- 18. 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;96:1086-1093.
- 19. Panait L, Bell RL, Roberts KE, Duffy AJ. Designing and validating a customized virtual reality-based laparoscopic skills curriculum. *J Surg Educ.* 2008;65:413-417.
- 20. Stassen LP, Bemelman WA, Meijerink J. Risks of minimally invasive surgery underestimated: a report of the Dutch Health Care Inspectorate. *Surg Endosc*. 2010;24:495-498.
- 21. Issenberg SB, McGaghie WC, Petrusa ER, Lee GD, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach*. 2005;27:10-28.
- 22. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigourous science for the assessment of surgical education and training. *Surg Endosc.* 2003;17:1525-1529.
- 23. Matell MS, Jacoby J. Is there an Optimal Number of Alternatives for Likert Scale Items? *Educational and Psychological Measurement*. 1971;31:657-674.
- 24. Madan AK, Frantzides CT. Substituting virtual reality trainers for inanimate box trainers does not decrease laparoscopic skills acquisition. *JSLS*. 2007;11:87-89.
- 25. Kanumuri P, Ganai S, Wohaibi EM, Bush RW, Grow DR, Seymour NE. Virtual reality and computer-enhanced training devices equally improve laparoscopic surgical skill in novices. *JSLS*. 2008;12:219-226.
- 26. Chmarra MK, Dankelman J, van den Dobbelsteen JJ, Jansen FW. Force feedback and basic laparoscopic skills. *Surg Endosc.* 2008;22:2140-2148.
- 27. Rosenthal R, Gantert WA, Hamel C, Metzger J, Kocher T, Vogelbach P, Demartines N, Hahnloser D. The future of patient safety: Surgical trainees accept virtual reality as a new training tool. *Patient Saf Surg.* 2008;2:16.
- Kossi J, Luorstarinen M. Virtual reality laparoscopic simulator as an aid in surgical resident education: two years' experience. *Scand J Surg.* 2009;98:48-54.
- 29. Van Dongen KW, van der Wal WA, Rinkes ICH, Schijven MP, Broeders IA. Virtual reality training for endoscopic surgery: voluntayr or obligatory? *Surg Endosc.* 2008;22:664-667.
- Verdaasdonk EG, Dankelman J, Schijven MP, Lange JF, Wentink M, Stassen LP. Serious gaming and voluntary laparoscopic skills training: a multicenter study. *Minim Invasive Ther Allied Technol.* 2009;18:232-238.
SAFETY AND COMPETENCE: A COMPARISON BETWEEN HEALTHCARE AND THE PETROCHEMICAL INDUSTRY

CHAPTER 3.1

Safety management, training and culture. What can we learn from the industry?

P.D. van Hove^{1,2}, R. van Wijk³, E.G.G. Verdaasdonk⁴, L.P.S. Stassen⁵, J. Dankelman¹

1: Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands 2: Department of Surgery, Reinier de Graaf Group, Delft, the Netherlands

3: Hydrocarbon management department, Royal Dutch Shell, Moerdijk, the Netherlands

4: Department of Surgery, Erasmus Medical Center, Rotterdam, the Netherlands

5: Department of Surgery, Maastricht University Medical Center, Maastricht, the Netherlands

ABSTRACT

Despite recent disasters, the petrochemical industry is known as a high-reliability sector, which has improved its safety performance in the last 20 years. Key elements in this improvement have been safety management systems, structured training and certification, and a high-level safety culture.

A safety management system is a structured way of managing all hazards that an organisation faces and includes inventory, analysis and management of risks. Along with improving safety these systems have made work processes in the petrochemical industry more efficient. These positive results led to the belief that other sectors could benefit from a same structured management approach. Currently, safety management systems are being implemented in all Dutch hospitals.

Training and certification are well structured in petrochemical industry. Every employee will eventually be trained and certified for general safety and specific tasks considering his working activities. This is repeated regularly. Healthcare has made major improvements in the field of training and certification, but compared to industry it still lacks compulsiveness.

The most important condition for establishing effective safety management is a high-level safety culture which is present throughout all layers of an organisation. A high-level safety culture facilitates open and blameless communication about risks and errors. Disasters in the oil industry have shown that a low-level safety culture can undermine the effect of a safety management system. Healthcare is currently characterised by a 'reactive' safety culture and will need to progress to a higher level to improve safety.

INTRODUCTION

In order to achieve a higher level of (patient) safety healthcare has been compared with other industries, especially aviation.¹⁻⁶ To a lesser extent comparisons have been made with the petrochemical industry, which is another high-risk industry with an effective safety policy. Although recent disasters such as the Prudhoe Bay (Alaska) oil spill in 2006 and the Deepwater Horizon oil spill in the Mexican Gulf in 2010 seem to contradict this, the implementation of "Safety Management Systems" (SMS) has given the industry insight in its risks and has improved safety over the last 20 years by adequate management of those risks.⁷ The recent accidents prove that slackening of attention for safety immediately increases risk.

It has been argued that an SMS can work similarly for healthcare⁸, which in The Netherlands has led to an advisory report for the Ministry of Health, written by the former CEO of Shell Netherlands.⁹ This report has led to the obliged implementation of SMS in all Dutch hospitals by 2013.

Is an SMS indeed an effective means to improve patient safety or is it just a theoretical solution? And will it stand on its own, or are certain preconditions required? The purpose of this paper is to give a clear description of an SMS and to emphasize the additional importance of training and establishment of a safety culture for the improvement of patient safety.

Petrochemical industry

The petrochemical industry is a high-risk industry with a history of numerous catastrophic accidents.⁷ A major turnaround in its safety management took place after the Piper Alpha disaster in the North Sea in 1988, in which human error was found to have caused gas leakage with a subsequent explosion and fire, costing the lives of 167 men.^{10,} ¹¹ It was this disaster which mainly led to the effectuation of the "Offshore Installations (safety case) Regulations" in the United Kingdom in 1993.¹² The Offshore Installations Regulations require petrochemical companies to have an effective SMS in function for their offshore installations. The same safety requirement is also found in the Seveso II Directive, which is legislation enforced by the European Union as early as 1976 and considerably revised in 1996.¹³

Despite these rigorous measures, several accidents have happened in the recent past. Nonetheless, numbers show that in general the implementation of structured SMS after the Piper Alpha disaster in 1988 has certainly caused a big leap towards improved safety performance in the petrochemical industry.⁷ This has eventually turned oil companies into high-reliability organisations (HRO). Although improving safety performance was the initial drive to implement SMS, companies noticed after implementation that reviewing and reorganising their processes in a structured way also made their workflow much more efficient and yielded them other (financial) benefits.

Safety Management Systems

A safety management system can be defined as "the systematic application of management processes to the problem of hazards an organisation faces".⁸ In his advisory report for the Dutch Ministry of Health⁹, CEO Willems stated four essential elements of a safety management system:

- 1. A system for risk inventory
- 2. A system for safe (blame-free) incident reporting
- 3. A method for analysis of incidents
- 4. A system to manage measures for improvement

As the name "safety management system" implies, it is a system which is implemented 'top-down' (Fig 1). However, this does not mean that all sorts of paternalistic rules are forced onto an organisation. The top-down structure essentially lies in the fact that the upper layers of an organisation initiate the safety policy and propagation thereof throughout the entire organisation, with the purpose to enforce ownership at every level. Therefore, the principle of an SMS is to set safety targets at the top, but to leave their realisation to the people doing the actual work. In this way all layers of an organisation have to participate actively in the safety policy and are not forced to passively follow orders.

The heart of an SMS is formed by the "Hazards and Effects Management Process" (HEMP). This structured process was developed at Shell in 1994¹⁴ and was based on Deming's famous management cycle "Plan – Do – Check – Act".¹⁵ The process starts with the identification of hazards within a process, and assessment of their possible consequences and their likelihood of occurrence. Then, plans and procedures are developed and implemented to reduce these effects to a minimum. The effects of the plans and procedures are monitored and might lead to adjustment when necessary (see Fig 1).¹⁶

The four essential elements of an SMS are needed for the functioning of a HEMP. A risk inventory and a system for (safe) incident reporting are used to populate registers of hazards and effects.¹⁴ Furthermore, a method for risk analysis is used to analyze the hazards and their effects. The "Bowtie" method is a tool which is frequently used in petrochemical companies to determine which barriers are needed to prevent incidents and which mitigations to reduce the consequences (see Fig 2). The bowtie analysis is a useful method to reveal at which locations in processes controls are needed.¹⁴



Figure 1 Structure of a safety management system with the hazards and effects management process at the centre. The safety management system consists of multiple layers, which have their own feedback mechanism and are interconnected. Based on Hudson et al.⁷



Figure 2 A bowtie diagram is centred around a potential undesired incident positioned in the middle. At the left-hand side are hazards which eventually can lead to consequences at the right-hand side when the incident in the middle has taken place. Barriers are incorporated in a process to prevent an incident. Mitigations are intended to reduce the consequences of the incident. Barriers and mitigations are called the 'controls' of the system.

Competency of personnel

In petrochemical industry proper training and certification are considered essential for safety. Not surprisingly regular training of all personnel for their tasks is common practice. This warrants that employees have knowledge about the processes they are working on and are skilled with the machinery required for their job.

Petrochemical industry has structured the training of personnel in separate phases. Before employment, general concepts of processes and equipment are taught in classrooms or simulation platforms. After the start of employment, every employee has to undergo several general safety trainings which have to be completed at a sufficient level before the employee is even allowed to enter the plant and start his actual duties. Further, more detailed, training is eventually most important and concerns the specific processes and/or specific hazards (e.g. dangerous substances) which the employee will encounter and are typical of his tasks. This training takes place by classroom learning, on-the-job training and, to a lesser extent, by the use of simulation. At the end of every training, employees are assessed mostly by taking both a theoretical and a practical exam. Only when these exams are finished with a sufficient result, an employee will be officially certified and allowed to perform the trained tasks. In order to make sure that all personnel is fully competent and maintains this status, all training is repeated on a regular basis. Some companies have a competency management system to monitor all training and certification of employees.

Commitment to training and certification is not always as strict as described here. It depends on the function an employee will fulfil. Eventually, all employees will undergo (safety) training, however for employees with an operational function safety is most urgent and therefore the situation described above does certainly apply to their situation. Possible consequences of ignorance were seen with the Deep Water Horizon disaster, where, apart from a lack of communication and bad management, inadequate training of engineering and rig personnel appeared to be an important contributing factor.¹⁷

Healthcare

Applying safety management to healthcare

A main difference between a safety management system in healthcare and other industries is that in other industries safety applies to a company's employees, whereas in healthcare safety applies to patients in the first place and possibly to employees in the second place. However, this does not change anything to the concept of a safety management system, which is to identify, assess and manage risks. To set up an SMS in healthcare the essential elements mentioned before should be installed in hospitals. This may sound straightforward, but is in fact a complex and time-consuming process, as became evident from the experience with the Dutch ambition to have an SMS implemented in all hospitals by 2008. This deadline was not met and had to be postponed until the end of 2012. Although some required components for an SMS were already present in hospitals, these were mostly not adequate. For example, incident reporting systems were present in most hospitals, but "safe" anonymous reporting was not possible.

To guide hospitals with all changes accompanying implementation of an SMS, a nationwide project was started in 2008 (VMS zorg: www.vmszorg.nl). The goal of this national project is to have every hospital manage its risks by implementation of a certified SMS and to have reduced preventable damage by 50% at the end of 2012. To help focus on the most important risks, ten evidence-based themes have been highlighted as 'high-risk themes' and for these themes concrete targets have been defined as well as interventions to meet these targets.

Training and certification

Together with patient safety, (skills) training has gained more attention in healthcare. However, healthcare has not yet succeeded in organising its training and certification as structured and formal as the petrochemical industry.

Education in healthcare can be roughly divided into two phases; basic education at medical university or nursery school and postgraduate education, which is eventually required for most functions in healthcare. Postgraduate education is perhaps the most important phase of education as it should provide the skills that are eventually needed for actual practice. Modern residency programs have improved, for instance by introducing the CanMEDS system of competency-based training.¹⁸ Nevertheless, one remaining fact is that the main part of postgraduate training exists of on-the-job training, which implies that errors are also being made on-the-job. Off-the-job training is practiced more and more but has brought along a timing problem. Due to busy schedules and clinical obligations it is often not possible to have trainees planned free from duties to allow them to attend courses at the desired moment. This is remarkable because required skills and knowledge will be lacking when needed, with loss of clinical effectiveness and potential risks as a result.

Besides these structural problems, content of training is also suboptimal. Healthcare workers are increasingly dealing with technology, but most training focuses on its professional application, whereas technical execution issues are ignored. For instance, in surgery, the main focus is on operative dexterity skills training, but how to actually use the increasingly complex equipment in the correct and safe way is usually not trained and might contribute to incidents with equipment.¹⁹ Likewise, non-technical (soft) skills like team-work, communication and attitude are just as important for improvement of safety and should receive more attention in training programs.^{20, 21}

Finally, formal assessment and certification are still in their infancy in healthcare.

For many courses, attendance is enough to receive a certificate as well as for accreditation for conferences. In this way, certification is in no way the proof of the actual possession of the required knowledge and/or skills. Fortunately, more formal and objective assessment is being introduced progressively in residency programs, although this process is still hampered by the lack of right methods for assessing the desired skills.²² Also, it may not always be possible to sharply define and demarcate competencies in healthcare, but this often is also the case in complex industries.

CULTURE

In petrochemical industry, the implementation of an SMS is what has eventually led to the evolution of its culture towards a high-level "safety culture". Safety is an issue within all layers of the organisation, whereas in healthcare it is still an occupation of enthusiastic individuals.

The safety culture within an organisation can be characterised by five levels, which were described by Hudson⁷ and are shown in Figure 3. The lowest level is the pathological culture where business has the highest priority, incidents are considered unimportant



Figure 3 The evolution of a safety culture. A pathological culture forms the lowest level and the highest level is called a generative culture. With ascending levels there is increased informedness and trust among employees. Adopted from Hudson et al.⁸

and caused by individuals. At the highest level is the generative culture where safety is integrated into all layers of an organisation and is a part of all work processes. Incidents are seen as failures of 'the system' and in such a culture one is continuously alert for possible threats, which can cause incidents. In between the pathological and generative culture are the reactive, calculative and proactive cultures. In calculative cultures, safety is primarily managed by the use of systems and is proclaimed by the management and not yet carried by its workforce. Along with the progression to each next level of a safety culture, an organisation becomes more informed by learning from its incidents, and there is increased trust among its workforce.⁷

At present, healthcare has reached the level of reacting; reacting to incidents, to legislation and to criticism from society. Systems for managing safety are being implemented, but are not fully active, so the calculative level is not yet reached. Therefore, healthcare is probably best characterized as reactive and busy becoming calculative. Nonetheless, as shown in Figure 3, to reach a true safety culture, evolution needs to progress further. Healthcare needs to learn to look and think ahead, which, in a safety perspective, is something uncommon and therefore it will cost great effort. It requires everybody at the work floor and in the management to put safety on his priority list and to accept addressing each other on safety issues to be the standard. It will also require to discuss errors openly without condemnation and to question one's own capabilities, to address shortcomings and to allow others to do the same. Only when these conditions are common practise, a higher safety culture will eventually be within reach.

Nevertheless, a safety culture remains a very dynamic phenomenon. The higher the level, the more effort it will cost to maintain that level. Fallbacks in the level of safety culture can undermine the effects of a safety management system, as was demonstrated with recent accidents. For example, a large explosion at an oil refinery in Texas City (2005) killing 15 employees was caused by overfilling and overheating of a raffinate splitting tower in combination with poor maintenance. It was a result of cost cutting and neglect of many safety regulations.²³ In 2006 in Prudhoe Bay in Alaska more than 800.000 litres of crude oil were spilled into the sea due to a leak in a pipeline and neglect of a leak detection system. The leak was caused by corrosion, due to cost cutting in maintenance.²⁴ A third example of inadequate priority of safety is the blowout of a drilling well in 2010, which led to the death of 11 people and a spill of 780 million litres of oil in the Mexican Gulf. In this case several decisions were taken which carried increased risk, but would cost less time and money than more safe alternatives. Company safety guidelines and risk analysis were ignored several times in these decisions.¹⁷

The common factor in these disasters is prioritising turnover over safety (pathological culture) in combination with poor leadership. These 2 factors subsequently led to the negligence and violation of safety rules and directives causing increased risks. Ironically,

the disasters have cost the companies much more money than cutting costs on safety would have ever rendered them.

These disasters are not exemplary for the industry as a whole, but testify of the essential importance of a safety culture as a precondition of any safety management: it will not work if it is not continuously supported by a safety culture, starting at the board level. This is equally applicable to healthcare, where safety is gaining popularity but still does not consistently have the highest priority.

CONCLUSION

Applying safety management systems as in the petrochemical industry, offers a possibility for healthcare to reduce accidents and improve patient safety. Training and certification are considered essential for safety in the petrochemical industry and could be better organised and more formal in healthcare. However, a change of culture is of most crucial importance to improve patient safety and to support a safety management system. To effectuate this, communication about errors, risks and safety will have to become much more open and less condemning among all layers in healthcare, so that lessons from accidents become consequently embedded in the process of care. Maintaining a high-level safety culture may even be harder, but is important as fallback could devastate the effect of a safety management system.

REFERENCES

- 1. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: cross sectional surveys. BMJ. 2000;320(7237):745-9.
- 2. Randell R. Medicine and aviation: a review of the comparison. Methods Inf Med. 2003;42(4):433-6.
- 3. Muller M. Safety lessons taken from the airlines. Br J Surg. 2004;91(4):393-4.
- Helmreich RL, Davies JM. Human factors in the operating room: interpersonal determinants of safety, efficiency and morale. In: Aitkenhead AA, editor. Bailliere's Clinical Anaesthesiology: Safety and risk management in anaesthesia. 2 ed. London, United Kingdom: Bailliere Tindall; 1996. p. 277-95.
- 5. Sax HC, Browne P, Mayewski RJ, Panzer RJ, Hittner KC, Burke RL, et al. Can aviation-based team training elicit sustainable behavioral change? Arch Surg. 2009;144(12):1133-7.
- 6. Helmreich RL. On error management: lessons from aviation. BMJ. 2000;320(7237):781-5.
- Hudson PTW. Safety Management and Safety Culture: The Long, Hard and Winding Road. In: Pearse W, Gallagher C, Bluff L, editors. Occupational Health and Safety Management Systems. Melbourne, Australia: Crowncontent; 2001. p. 3-32.
- Hudson P. Applying the lessons of high risk industries to health care. Qual Saf Health Care. 2003;12 Suppl 1(i7-12.
- 9. Willems R. Hier werk je veilig, of je werkt hier niet. The Hague: Shell Netherlands. 2004.
- 10. Cullen The Lord. The Public Enquiry into the Piper Alpha Disaster. London, United Kingdom. 1990.
- 11. Pate-Cornell ME. Learning from the Piper Alpha Accident: A Postmortem Analysis of Technical and Organizational Factors. Risk Analysis. 1992;13(2):215-32.
- 12. Flin RH, Slaven GM. The selection and training of offshore installation managers for crisis management. London, United Kingdom: HSE Books. 1994.
- 13. European Council. Council Directive 96/82/EC of 9 December 1996 on the control of majoraccident hazards involving dangerous substances. 1996
- 14. Salter N, Siddoway M. Implementation of the Hazards and Effects Management Process (HEMP). AIChE Spring National Meeting; New Orleans, LA2004. p. 1064-75.
- 15. Deming WE. Out of the Crisis. Cambridge, Massachusetts: Massachusetts Institute of Technology, Center for Advanced Educational Services; 1982.
- 16. Chamberlain GA. Management of large LNG hazards. 23rd World Gas Conference; Amsterdam, The Netherlands2006. p. 2474-84.
- B. Graham, WK. Reilly, F. Beinecke, DF. Boesch, TD. Garcia, CA. Murray, et al. Deep Water. The Gulf Oil Disaster and the Future of Offshore Drilling. Report to the President. Washington DC. 2011 www.oilspillcommission.gov.
- 18. Frank JR. The CanMEDS 2005 Physician Competency Framework. Better standards. Better physicians. Better care. Ottawa, Canada. 2005.
- 19. Verdaasdonk EG, Stassen LP, van der Elst M, Karsten TM, Dankelman J. Problems with technical equipment during laparoscopic surgery. An observational study. Surg Endosc. 2007;21(2):275-9.
- 20. Helmreich RL, Schaefer H-G. Team performance in the operating room. In: Bogner MS, editor. Human error in medicine. Hillsdale, NJ: Lawrence Erlbaum Associates; 1994. p. 225-53.
- 21. Lange JF, Dekker-van Doorn CM, Haerkens MH, Klein J. [A safety culture in hospitals.]. Ned Tijdschr Geneeskd. 2011;155(18):A2253.
- 22. Van Hove PD, Tuijthof GJM, Verdaasdonk EGG, Stassen LPS, Dankelman J. Objective assessment of technical surgical skills. British Journal of Surgery. 2010;97(7):972-87.

- 23. U.S. Chemical Safety and Hazard Investigation Board. Investigation Report Refinery Explosion and Fire. Texas City, Texas. 2007. p. 1-341.
- 24. B. Chapman, J. Alkire, B.J. Harris, B. Vest, R.G. Buckendorf, S. Barnes, et al. GC-2 Transit Line Spill Prudhoe Bay Western Operating Area March 2, 2006. Incident investigation report. 2006.

CHAPTER 3.2

Training and assessment of equipmentrelated competence. Comparison of a petrochemical company and a hospital

P.D. van Hove^{1,2}, R. van Wijk³, E.G.G. Verdaasdonk⁴, L.P.S. Stassen⁵, J. Dankelman¹

1: Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands 2: Department of Surgery, Reinier de Graaf Group, Delft, the Netherlands 3: Hydrocarbon management department, Royal Dutch Shell, Moerdijk, the Netherlands 4: Department of Surgery, Erasmus Medical Center, Rotterdam, the Netherlands 5: Department of Surgery, Maastricht University Medical Center, Maastricht, the Netherlands E-pub ahead of print: Health Technol. 2013 DOI 10.1007/s12553-013-0041-z

ABSTRACT

The role of technology in healthcare is constantly increasing. Training with equipment in healthcare does not seem to be as strictly organized as in other industries. The purpose of this paper is to describe the differences in organization and management of equipment-related training between a petrochemical company and a hospital, in order to extract information that can be incorporated into a hospital's safety management system.

Interviews were held at a petrochemical plant of one of the world's largest companies and at a large non-academic teaching hospital, both in The Netherlands. At the petrochemical company, 2 learning advisors were interviewed, responsible for the content of training programs for operators. At the hospital, interviews were held with 2 educational coordinators and a team leader of the operating department.

At the petrochemical company competencies are defined for every job including accompanying training sessions and assessments, followed by certification. All training and certification is mandatory. There is a separate department responsible for all education within the company and all competencies and certification are registered electronically. At the hospital training is not organized as strictly. Equipment-related training is not mandatory and, except for laparoscopic operating equipment, not repeated regularly. Assessment and certification of personnel does not take place.

Equipment-related training and assessment in hospitals could be organized more strictly. The petrochemical industry can be used as an example of a sector in which this development has already taken place. To establish such changes, sharpening of legislation and regulations by governmental bodies is of fundamental importance.

INTRODUCTION

In past years, the amount of equipment used in healthcare has increased enormously, especially in surgical specialties. This increase of technology has made many procedures more complex and has led to more training initiatives to master this complexity. This training however, focuses largely on dexterity whereas handling the equipment is ignored.¹ Nonetheless, it has been shown that incidents with equipment are not infrequent.²⁻⁴

An industry which has made major improvements in safety and training policies is the petrochemical industry. Many methods for incident analysis, like TRIPOD and PRISMA, originating from aviation, were successfully used in the petrochemical industry and are now being used in hospitals as well. By the end of 2012 all Dutch hospitals are expected to have implemented a safety management system. In the petrochemical industry an essential factor related to safety is the use of equipment. History has taught that mistakes can lead to the death of many people.⁵ It is therefore that this industry focuses heavily on proper training of its employees.⁶ Furthermore, as many different types of equipment are used in different processes, different skills are required in different jobs. Therefore, when considering management and organization of equipment-related education, a petrochemical plant has more similarities with a hospital than, for instance, aviation does. For this reason, in this study a comparison is made between a petrochemical plant and a hospital.

The purpose of this study is to compare the management of equipment-related competence between a petrochemical company and a hospital, in order to identify important differences and to give recommendations to improve management of equipmentrelated education and competence within healthcare institutions.

METHODS AND MATERIALS

Interviews have been held at a large multinational petrochemical company and at a large non-academic teaching hospital, both in The Netherlands. At the petrochemical company two 'learning advisors' from the department for 'learning and development' have been interviewed. At the hospital, a coordinator for 'medical educational programs', a coordinator for the 'knowledge and information centre' (including a large skills lab) and a team leader of the operating department have been interviewed.

The petrochemical company studied is one of the world's largest multinationals in the petrochemical industry. The interview took place at one of its largest plants. This plant has two locations close to each other which have a combined 'learning and development' department. This department was founded in 2009 and is responsible for all training and education of about 2600 employees working at both locations. The learning and development department is led by a learning manager who supervises three 'learning advisors'. Of these three learning advisors, two are responsible for the learning and development policy at the separate plants and one is responsible for an educational program for all operators at both plants. Operators fulfill one of the most essential functions at a plant, i.e. an operator manages all processes and makes sure all machinery is working properly in order to ensure safe and optimal production. He is therefore directly responsible for the core-business of the plant and can be compared with medical staff in a hospital.

The hospital studied is a large non-academic teaching hospital with 3000 employees working on one location with more than 600 beds. In this hospital a special department for education exists since 2010. At the moment of the interview this department was still in a process of expanding its position. Its current focus is mainly on a general introductory program for new employees and yearly resuscitation training sessions for all personnel. Equipment handling is not yet part of any training featured or managed by this department.

In both interviews several subjects with regard to training and competence management related to equipment handling have been discussed. In the results section a comparison is made of the organization of equipment-related training and assessment within these two institutions. Table 1 shows the most important differences. After the discussion, conclusions and recommendations are given.

RESULTS

Petrochemical company

At the petrochemical company, the management is responsible for setting training standards and for providing structure and facilitation for execution of the required training programs. This execution is done by the learning and development department. The training and certification of all employees is registered in order to keep an overview of all competencies within the company and separate departments. This is done with an electronic learning management system. To warrant proper training at all levels within a plant, there are three company-wide educational programs. Equipment-related training is embedded in all these programs, which will be described here.

Basic technical and safety training for all employees

It is mandatory within the petrochemical industry that every employee possesses a basic level of technical and safety knowledge and skills (http://osha.europa.eu). These basic competencies ensure that, in case of an emergency, all employees are familiar with

| | Petrochemical company | Hospital |
|--|-----------------------|-----------------------------|
| Equipment training | | |
| For all involved employees | Yes | No, only OR nurses |
| New employees | Yes | No |
| Present employees | Yes | Only laparoscopic equipment |
| New equipment | Yes | Yes |
| Present equipment | Yes | Only laparoscopic equipment |
| Repetition of training | Yes | Only laparoscopic equipment |
| Mandatory | Yes | No |
| Assessment | Yes | No |
| Registration | Yes | Yes |
| Certification | Yes | No |
| Educational department | Yes | Yes |
| LMS* | Yes | Partly |
| All education centralized | Yes | No |
| Defined competencies for every function | Yes | No |

Table 1Most important differences in organization of equipment-related training and assessmentbetween a petrochemical company and a hospital.

* LMS: learning management system

the essential procedures and skills. These procedures and skills vary from knowledge of extinguishing a fire to the ability of safely shutting down a whole production unit of the factory. Due to this variation in skills, the required basic technical and safety competencies vary among the different functions and staff within a plant. The basic safety training program ensures that every new employee at a plant is provided with the proper training concerning the knowledge and skills belonging to his/her function. This includes the ability to operate the equipment needed during an emergency. After the training, the employee has to take an assessment to warrant that he is competent.

Educational program for operators in training

The educational program for operators is a full traineeship, in which operators are educated from a basic theoretical pre-employment level to a fully-specialized operator. The program consists of a general part and a specialization part. In the general part young operators learn to apply basic theoretical knowledge in practice. In the specialization part an operator focuses on a specific unit of the plant he is working at. This consists of a theoretical part (class room, or e-learning) and an on-the-job training, guided by more experienced operators. Depending on the kind of specialization, the whole trajectory takes 4 to 9 years. The operators' educational program is structured into multiple promotional levels. For every level a set of competencies has been defined, including equipment-related skills. All competencies are measurable and assessable. In order to be promoted to a next level, the operator-trainee has to pass an assessment of all the required competencies for that level. This may take up to 3 hours and usually consists of an oral exam in combination with a practical assessment at the work floor or in a process simulator. Because promotion of a trainee depends of the assessment result, the assessments are always taken by two assessors. These assessors are usually department supervisors who have had a special training.

Educational program for established operators

This program has been developed to warrant competence among established operators as well. Like with the program for operators in training, measurable and assessable competencies have been defined for this program as well. These are not the same as for the trainees, but a selection has been made of certain 'critical competencies' an operator should possess. A competency is defined as 'critical' when it is of essential importance in the work of an operator. Usually these competencies are procedural equipment-related. An example of a critical competency is: 'safe shutdown of electrical equipment'. This is an important procedure when for instance maintenance has to take place on electrically powered equipment. If not done properly, employees performing the maintenance work are at risk of injury caused by machinery. Part of the assessment of this competency is for an operator to show in practice that he is capable of shutting down and isolating the machinery safely.

Assessment of critical competencies for established operators is repeated every 3 years and is performed by specially trained supervisors. As with the operators' educational program, an assessment exists of an oral exam and a practical assessment at the work floor or in a process simulator. To pass the assessment the operators have to score at least 75% and for some specific competencies even 100%. In total, they can take 3 attempts for every assessment, but nearly all operators pass with their first attempt.

To prepare for their assessments the operators can use e-learning modules and the protocols of their plant-unit, but there is no mandatory study-material. Classroom learning is deliberately avoided, because practice has shown, that it has a low efficiency and it is difficult to organize combined training sessions for operators on shift.

Introduction of new equipment

When new equipment is being introduced to a plant a special training and assessment cycle is designed and set up to train and assess all employees who will be using the new equipment. Where possible, simulation is being used for this purpose as it offers the opportunity to practice multiple scenarios in a production process without interfering with the real process.^{7,8}

Hospital

In the hospital, just like at the petrochemical plant, the management is responsible for defining goals and criteria for education. The educational department is only partly responsible for execution hereof, as many departments have their own responsibilities regarding national residency (or nursing) training programs.

Within the educational department the main focus is on education which the hospital is required to facilitate by the healthcare inspectorate and other certification bodies. This mainly consists of a general introductory course for new employees, which focuses on general safety issues and introductions to supporting services like computer programs which have to be used during the daily work. Equipment handling per se is not part of this course or any of the other courses organized by the educational department. This kind of education is organized by departments themselves. Therefore, an additional interview was held with a team leader at the operating department to discuss the organization of equipment training at that department.

Equipment-related education at the operating department

At this department training is focused on the operating room (OR) nurses. Medical specialists and other personnel are welcome to join, but their rate of attendance is often low.

Training with equipment is organized following the introduction of new equipment. Multiple sessions are organized in which a representative from the production company gives an explanation and demonstration of the equipment. These sessions are not obligatory, but are usually attended by the majority of the OR nurses. Attendance is registered at the department, but there are no consequences for not attending. A problem with attendance usually is the limited number of sessions in combination with the rotational schedules of the OR nurses. OR nurses who have not attended a training session usually get an explanation from a colleague at another time.

Laparoscopic equipment is treated as an exceptional group of equipment for which the Dutch Healthcare Inspectorate (DHI) has put up several requirements. One of these demands is that the OR-personnel is trained regularly in the use of this equipment.⁹ In this hospital therefore, a yearly training cycle has been set up especially for laparoscopic equipment. These training sessions consist of a short demonstration of the equipment, but the focus is on scenario-based sessions in small groups. During the training, problems with the equipment are simulated and the nurses are challenged to solve these problems. Like training with new equipment, these training sessions are not obligatory either.

Both types of training are not combined with any form of assessment or certification as a guarantee of competence. Registration of attendance takes place for use during inspections of the DHI. An equipment-related training program for new employees does not exist, so officially new employees do not receive any equipment-related training at the start of their job. They have to wait for this until the next yearly cycle for laparoscopic equipment takes place. For other equipment they have to depend on informal instructions by colleagues.

DISCUSSION

This paper gives a description of the organization and management of equipmentrelated training in two institutions from two different industries. Both industries are high-risk industries with a large influence of technology. As was shown in this paper, there is a clear difference in the way these institutions train their personnel with regard to technology. Whereas the studied petrochemical company has a very strict educational program with a clear definition of the required competencies for every function within the company, this education is less strict and much more open-ended in the studied hospital.

The authors would like to emphasize that this paper addresses equipment-related training and education and not education of personnel in general. Secondly, for this paper, two institutions were interviewed and it could be argued whether these institutions are representative for their sector. However, it was not the purpose of this paper to give a precisely weighted representation of the average in both sectors, but to give an impression of the differences in the way that personnel is trained for equipment handling in their job. Thus, it is not the intention of the authors to state that the situation as described for the studied organizations is typical of other organizations in the same sectors. Nonetheless, the petrochemical company is one of the world's largest oil companies and therefore has an exemplary function in its sector. The hospital is a large non-academic teaching hospital and has a separate department for education, which is not common for a non-academic hospital. Teaching hospitals have to comply with all sorts of rules and regulations, especially concerning education. For these reasons it is not unlikely that both institutions are qualitatively better than the average in their sector.

One of the most notable differences is the strict organization of training within the petrochemical company compared to the open-ended organization in the hospital. The strict organization of education within the petrochemical company is based on legislation for the petrochemical industry concerning this topic. Due to multiple severe disasters in the 1970's and 1980's the "Offshore Installations Regulations" were instituted in the United Kingdom in 1993. These regulations were based on a detailed report containing numerous recommendations regarding training, assessment and certification of offshore installation managers.⁶ In healthcare such detailed legislation does not exist, but general demands have been put up in The Netherlands by the DHI.⁹ On a European level, the European parliament is preparing a law on patient safety in which training with

equipment will be mentioned.¹⁰ This paper illustrates that as long as no strict demands are proclaimed by governments, strict organization of equipment-related education and certification of personnel is probably not prioritized by healthcare institutions. This statement is strengthened by the less open-ended organization of training with laparoscopic operation equipment in the hospital. The special requirements of the DHI for this group of equipment have led to intensified training. Especially the scenario-based training sessions in small groups refer to practical training sessions at the petrochemical company. However, assessments and certification are still lacking. Therefore, the basis for a strict organization of training and certification of healthcare personnel is legislation, enforced by healthcare inspectorates. Such legislation should provide healthcare institutions with requirements for content and frequency of training and a demand for assessment of competence. Secondly, there should be 'ownership' and facilitation of a strict organization by the leadership of the hospital.

Another difference between the petrochemical plant and the hospital is the periodical training and assessment of established operators compared to no equipment-related training and assessment at all for medical specialists. OR nurses do train yearly, but only for the handling of laparoscopic equipment and not for other equipment. It is debatable which function can be best compared with the function of operator, as, in practice, the OR nurse and the medical specialist are shared users of the equipment. However, the specialist is mainly concerned about the result or effect that is needed from the equipment, whereas the OR nurse is usually the one connecting and setting the device. In this way, the medical specialist could more or less be defined as the supervisor of the OR nurse. Nonetheless, there is a certain degree of overlap between both functions which consequently translates into the required skills and knowledge for proper use of the equipment. Finally, the medical specialist is legally responsible for the whole operation, including the equipment employed.¹¹ Therefore, to warrant safety, both the OR nurse and the medical specialist will need to be trained for equipment handling, although the focus of these trainings can differ for both functions.

Due to the previous, the impression might have occurred that healthcare functions under all sorts of loose standards. However, nowadays, quality of care and patient safety are taken very seriously, which is illustrated by an increasing number of standards and the many bodies and organizations which check the compliance of healthcare institutions to these standards. For instance, one of the largest international accreditation organizations is the Joint Commission International (JCI). This organization accredits healthcare institutions for quality and patient safety and has become leading worldwide. Their standards also include demands for training and certification of personnel.¹² However, these standards form a framework and do not prescribe detailed instructions for the exact subjects of training. Nonetheless, accreditation bodies as the JCI are crucial for maintaining high standards of quality and safety in healthcare. Finally, there are

certainly some medical specialties which have been taking action on the field of training with equipment. In anesthesiology, for instance, training with equipment has been common for some years.¹³

It could be argued whether it is still realistic to expect the OR nurses and medical specialists to have detailed knowledge of all equipment they use. Medical technology is expanding rapidly and getting more and more complex, which asks for an increasing effort of its users to keep up with developments. Additionally, other responsibilities (e.g. administrative, educational, legal) are expanding as well. Considering this, it could be considered whether operating equipment should still be one of those responsibilities. A solution could be to, fully or in part, transfer the responsibility for the equipment to an 'OR equipment specialist' (i.e. a clinical physicist or an equipment engineer) in the future. However, in the mean time OR nurses and medical specialists should be able to use their equipment safely and therefore training seems inherent for both.

To identify the exact content for training, establishment of a framework of competencies required for every separate function, as was done in the petrochemical company, can serve as a guideline. This framework should then be the basis from which the content and frequency of training and assessment can be made up. Eventually, every assessment should lead to certification, which could even be used as a condition for prolongation of ones employment contract. In that way, the educational policy is best illustrated by a phrase once used by a former CEO of the exemplified petrochemical company: "you work here safely, or you don't work here at all".

CONCLUSIONS AND RECOMMENDATIONS

Compared to a petrochemical company, equipment-related training in a hospital is open-ended. There is no regular repetition of training, no assessment of competence and no certification. Laparoscopic equipment seems to be an exception with regard to regular repetition. Adequate regulation and legislation is of fundamental importance to establish stricter organization and management of training and thereby patient safety.

Following these conclusions, a number of recommendations can be made:

- Address the role of hospital boards in proclaiming and facilitating stricter organization of training of equipment handling
- Establish required competencies for every job
- Make training with equipment mandatory for its users
- Set and execute training cycles, including repetition
- Set and execute assessment cycles
- Set up a system for certification

REFERENCES

- 1. van Hove PD, Verdaasdonk EG, van der Harst E, Jansen FW, Dankelman J, Stassen LP. Effect of Basic Laparoscopic Skills Courses on Essential Knowledge of Equipment. Surg Innov. 2012.
- 2. Verdaasdonk EG, Stassen LP, van der Elst M, Karsten TM, Dankelman J. Problems with technical equipment during laparoscopic surgery. An observational study. Surg Endosc. 2007;21(2):275-9.
- 3. Courdier S, Garbin O, Hummel M, Thoma V, Ball E, Favre R, et al. Equipment failure: causes and consequences in endoscopic gynecologic surgery. J Minim Invasive Gynecol. 2009;16(1):28-33.
- 4. Panesar SS, Salvilla SA, Patel B, Donaldson SL. Laparoscopic cholecystectomy: device-related errors revealed through a national database. Expert Rev Med Devices. 2011;8(5):555-60.
- Pate-Cornell ME. Learning from the Piper Alpha Accident: A Postmortem Analysis of Technical and Organizational Factors. Risk Analysis. 1992;13(2):215-32.
- 6. Flin RH, Slaven GM. The selection and training of offshore installation managers for crisis management. London, United Kingdom: HSE Books. 1994.
- 7. Merrit R. Process simulators aren't just for training. Control; April 2006. 2006.
- 8. McArdle S, Cameron D, Meyer K, editors. The life cycle simulator: From concept to commissioning... and beyond. SPE Intelligent Energy Conference and Exhibition. 2010;1:611-20
- 9. Inspectie voor de Gezondheidszorg (IGZ). Risico's van medische technologie onderschat. Den Haag: Inspectie voor de Gezondheidszorg (IGZ). 2008. p. 1-81.
- 10. European Council. Council recommendation on patient safety, including the prevention and control of healthcare associated infections. 2009
- 11. Orde van Medisch Specialisten. Verantwoordelijkheid medisch specialist bij onderhoud en beheer van medische apparatuur. Leidraad. Utrecht. 2008.
- Joint Commission International. Joint Commission International Accreditation Standards for Hospitals. 4th Edition. Standard lists versions. Illinois: Joint Commission International; 2011. p. 1-40.
- 13. Waldrop WB, Murray DJ, Boulet JR, Kras JF. Management of anesthesia equipment failure: a simulation-based resident skill assessment. Anesth Analg. 2009;109(2):426-33.

TRAINING FOR SAFE HANDLING OF TECHNICAL LAPAROSCOPIC EQUIPMENT

CHAPTER 4.1

Use of checklists for laparoscopic equipment in Dutch hospitals

P.D. van Hove^{1,2}, E.G.G. Verdaasdonk³, J. Dankelman¹, L.P.S. Stassen⁴

1: Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands 2: Department of Surgery, Reinier de Graaf Group, Delft, the Netherlands 3: Department of Surgery, Erasmus Medical Center, Rotterdam, the Netherlands 4: Department of Surgery, Maastricht University Medical Center, Maastricht, the Netherlands

Published in: Ned Tijdschr Geneeskd. 2012; 156(33):A4505, in Dutch

ABSTRACT

Objective A survey among Dutch hospitals about the use of preoperative checklists and training concerning operating equipment for laparoscopic surgery.

Design Telephone-based questionnaire.

Methods Between June and August 2011 all Dutch hospitals were contacted by phone to complete a questionnaire consisting of 19 questions.

Results In total 89 of 91 hospitals completed the questionnaire: 8 academic hospitals, 44 teaching hospitals and 37 general hospitals. Fifty hospitals (56.2%) used a checklist for laparoscopic operating equipment. In the remaining 39 hospitals (43.8%) no specific checklist was used, but a standard procedural pre-use check of the equipment was performed. The level of detail differed among the different checklists. On 9 of the 50 checklists different devices were not addressed separately and on half of all checklists connections were not a check-item.

In 87 hospitals (97.8%) personnel was trained in the use of operating equipment. In 33 hospitals (37.1%) this only took place with the introduction of new equipment, but in other hospitals training took place more often. In 58 hospitals (66.7%) training with equipment was mandatory.

Conclusion The use of detailed preoperative checklists for the use of laparoscopic operating equipment is not yet the standard in all Dutch hospitals. A procedural pre-use check is performed in all hospitals, but not everywhere this is formally documented with the use of a checklist. In one third of the surveyed hospitals training with operating equipment was not mandatory.

INTRODUCTION

The practice of minimally invasive surgery of the abdomen, laparoscopic surgery, in Dutch hospitals has increased from the '90s. In 2009 more than 37.000 laparoscopic operations were performed. In 2007 this number was around 33.000 (http://statline. cbs.nl/statweb/). A report of the Dutch Healthcare Inspectorate (DHI) from the same year, entitled "Risks of minimally invasive surgery underestimated", stated that severe incidents had occurred regularly during minimally invasive operations.¹ Observational studies have indeed revealed that in 42 – 87% of laparoscopic operations incidents occur with the equipment used. These incidents differ from equipment not being present to faulty connections of equipment or its accessories.^{2, 3} Initially, this sort of incidents leads to a delay and attracts the attention.^{2,4} However, the Dutch institute for Healthcare Research, NIVEL, calculated that in 5% of all laparoscopic operations injury occurs as a consequence of the use of medical technology. Mentioned explicitly were: anastomotic leakage, infections and permanent injury of gall tract, ureter and caecum.⁵

According to the DHI the leading causes of incidents were a lack of competence among users of the new technology and insufficient standardization of moments of (patient) information transfer in the operative process.^{1, 6} The latter not only referred to equipment, but primarily to the operative process in total with all its information transfers. In a reaction to the findings in its report, the DHI advised the implementation of checklists to structure moments of information transfer. Furthermore, implementation of a time-out procedure, as a final preoperative check, was enforced.⁷ During a time-out procedure the entire operating team discusses which patient is to be operated on, which operation will be performed and if there are any special considerations such as allergies. Another point of attention during this procedure is whether all required persons and materials are present. The specific focus on equipment during the time-out procedure is limited, as the presence, positioning, settings and connections are no official check items. However, deficiencies regarding these specific aspects have been reported.²

It is unknown in how many Dutch hospitals a structured checklist is currently being used for the preoperative checking of laparoscopic operating equipment and how detailed these checklists are. The frequency of training with the equipment is unknown as well. The purpose of this study is to regain insight into the current status regarding these two issues of safety concerning laparoscopic operating equipment in Dutch hospitals.

METHODS

Between June and August 2011, 91 Dutch hospitals were contacted by telephone to answer a questionnaire on the use of checklists and performance of training concerning

laparoscopic operating equipment. In this study the term 'laparoscopic operating equipment' includes the equipment which is always used during laparoscopic operations, such as an insufflator, a light source, a camera unit and a monitor. It also includes other equipment which is used both during laparoscopic operations and open operations, such as suction equipment and electrosurgical equipment. Laparoscopic instruments, such as scissors and graspers are outside the scope of this study.

Hospitals which did not reply initially were contacted repeatedly to get a good response rate. Contact was sought with the operating department of the hospitals, preferably with the team leader, an operating nurse specialized in laparoscopic surgery or another person who was responsible for the operating equipment and had knowledge on the use of checklists and organization of training. This person was asked to answer a short questionnaire of 19 questions in total (Table 1). All answers were entered into a database. After completion of all questionnaires and before analysis all data were made anonymous. Data were analyzed with SPSS for Windows, version 16.0.

RESULTS

Checklists

Of all 91 hospitals 89 (97.8%) participated in the study by answering the questionnaire. One hospital explicitly did not want to cooperate due to a lack of time and one hospital repeatedly did not reply to the request for participation.

The use of checklists among different types of hospitals is presented in Table 2. The mean time that a checklist was in use was 20 months (2 weeks – 48 months). To get an idea of the level of detail of the checklists, it was inventoried whether insufflator, light and imaging equipment, electrosurgical equipment and other equipment were separately addressed on the checklists and what were the specific checkpoints. These results are presented in Table 3. Table 4 gives an overview of the differences in timing of the checking procedures.

In 39 of the 50 hospitals where a checklist was used (78%), this checklist was selfdesigned. In eight hospitals (16.0%) a pre-existent checklist or a slight modification was used. In 3 hospitals the origin of the checklist was unknown to the interviewed person.

In 6 of the 50 hospitals where a checklist was used (12.0%), this was an electronic version. In the other 44 hospitals (88.0%) the checklists were paper-based, but it was often stated that design of an electronic version was in progress.

| Table 1 Questionnaire used in Telephone-based inter | view |
|---|------|
|---|------|

Questions

Use of checklist and checklist characteristics

- 1 Is a preoperative checklist being used, which is specifically aimed at laparoscopic operating equipment or which is wider orientated but with a specific part focused on laparoscopic equipment?
- 2 If no, does preoperative checking of this equipment take place without the use of a checklist?
- 3 If a checklist is being used: how long is it in use?
- 4 Is the checklist self designed or was it pre-existent?
- 5 Is the checklist electronical or paper-based?
- 6 When is the checking procedure (with or without checklist) performed?

Which of the following checkpoints are part of the checklist?

- 7 Presence of the equipment
- 8 Completeness of the equipment
- 9 Functioning / defects of the equipment
- 10 Connections

Which of the following devices are addressed separately on the checklist?

- 11 Insufflator
- 12 Light and imaging equipment
- 13 Electrosurgical equipment
- 14 Other equipment (OR-lamps, OR-table, suction equipment etc.)

Training

- 15 Does training with equipment take place and who participate in it?
- 16 Does training take place regularly or irregularly?
- 17 How often does training take place on average?
- 18 Which training method is being used?
- 19 Is training mandatory and if yes, for who is it mandatory?

 Table 2
 Number of hospitals in which a checklist is used for preoperative checking of laparoscopic operating equipment.

| Type of hospital | | Checklist | No checklist |
|--------------------|----|-----------|--------------|
| | N | N | N |
| All | 89 | 50 | 39 |
| Academic hospitals | 8 | 3 | 5 |
| Teaching hospitals | 44 | 26 | 18 |
| General hospitals | 37 | 21 | 16 |

| _ | • | 1 1 | 5 |
|-----------------------------|--|------------|---------------|
| La | paroscopic equipment | Checkpoint | No checkpoint |
| In | sufflator | 40 | 10 |
| - | presence | 39 | 11 |
| - | completeness | 38 | 12 |
| - | functioning / defects | 40 | 10 |
| - | connections | 24 | 26 |
| Light and imaging equipment | | 40 | 10 |
| - | presence | 39 | 11 |
| - | completeness | 38 | 12 |
| - | functioning / defects | 40 | 10 |
| - | connections | 24 | 26 |
| Ele | ectrosurgical equipment | 39 | 11 |
| - | presence | 38 | 12 |
| - | completeness | 37 | 13 |
| - | functioning / defects | 39 | 11 |
| - | connections | 23 | 27 |
| Ot | her equipment (OR-lamps, OR-table, suction | 33 | 17 |
| eq | uipment etc.) | | |
| - | presence | 33 | 17 |
| - | completeness | 32 | 18 |
| - | functioning / defects | 33 | 17 |
| - | connections | 20 | 30 |
| Eq | uipment not mentioned explicitly | 9 | |
| - | presence | 9 | |
| - | completeness | 8 | |
| - | functioning / defects | 8 | |
| - | connections | 0 | |

Table 3Separately addressed devices and checkpoints on the checklists.

The numbers concern the 50 hospitals where a checklist for equipment is being used.

 Table 4
 Timing of the preoperative equipment check, either with or without the use of a checklist.

| Timing of the check | Checklist | Procedure | Total |
|------------------------|-----------|-----------|-------|
| Beginning of the day | 18 | 11 | 29 |
| Before the first use | 11 | 8 | 19 |
| Before every operation | 21 | 20 | 41 |
| Total | 50 | 39 | 89 |

Training

In 87 of the 89 participating hospitals training with operating equipment was performed. Figure 1 gives an overview of different training methods and their prevalence within the interviewed hospitals.

In 33 hospitals (37.9%) training only took place with the introduction of new equipment. In the other 54 hospitals the interviewed employee described the frequency of training as 'regular' in 43 hospitals and as 'irregular' in 11 hospitals. Figure 2 gives an



Figure 1 Prevalence of application of different training methods. Some hospitals used multiple methods which causes the total to exceed 87.



Figure 2 Frequency of training with equipment.

overview of the frequency of training with equipment. Training with equipment was mandatory in 58 hospitals (66.7%). However, in 15 hospitals (17.2%) this obligation was limited to the operating room nurses who were on duty at the moment of training. In 29 hospitals (33.3%) training with equipment was not mandatory. In almost all hospitals the training sessions were aimed at operating room nurses. Only 4 hospitals indicated that their training sessions were also specifically aimed at residents and medical specialists.

DISCUSSION

The primary focus of this study was the preoperative checking of laparoscopic operating equipment. This is performed in all 89 surveyed hospitals and in 50 hospitals (56%) this is performed by the use of a preoperative checklist. The difference in level of detail among the different checklists was remarkable.

This study was performed by the use of a telephone-based survey. Preferably the questionnaire was answered by a team leader of the operating department or a specialized operating room nurse. In all hospitals the interviewed employee was was deemed to have most knowledge on the use of checklists within the hospital. The response rate in this study was high (98%). However, the precision of the obtained information depends on the knowledge of the interviewed employee and this was not verified by interviewing of additional employees within the same institution. This is a limitation of this study.

It is not easy to determine the optimal level of detail for a checklist. A checklist with too much detail will cause delay and irritation among its users, which may lead to 'checklist fatigue' whereby the checklist is experienced as a burden and not as an aid. A good checklist should only contain the critical steps of a procedure.⁸⁻¹¹ A step is defined as 'critical' when its omission will have serious consequences.¹¹ Nonetheless, it remains debatable what exactly are the critical steps of a procedure, because seemingly unimportant events can eventually have major consequences.¹² With regard to operating equipment the Dutch guidelines for the preoperative and peroperative trajectory prescribe that during the time-out procedure it should at least be checked whether the equipment 'is present and checked.'^{13, 14} This certainly is a critical step, but perhaps its formulation is too vague to prevent errors with the equipment. It would be more explicit to describe several checkpoints for every separate device.¹⁵ A study with a checklist which was composed in such manner showed to reduce the number of incidents with equipment with 53%. On that checklist all separate devices and their accompanying connections and settings were addressed separately.¹⁶

It is debatable whether another checklist added to the existing number of checklists will not overshoot its target. The use of multiple checklists, which all focus in detail on separate steps of a whole process, might inflict on the overview of a process and with
that on its effectiveness. To maintain overview of the operating process the time-out procedure is explicitly suitable. In the specific case of equipment checking, an OR nurse could complete a detailed checklist before the time-out procedure takes place. During the time-out procedure a short mentioning of 'equipment present and checked' then refers back to this checklist. In this way a detailed check of the equipment takes place and effectiveness and overview of the whole operating process is maintained.

Preoperative checking of equipment does not stand by itself. Maintenance of the equipment and training of its users is equally important in preventing incidents. According to the Dutch guidelines on the pre- and per-operative process, healthcare institutions should warrant documentation of procedures considering the whole process of purchase, introduction, maintenance and use of equipment.^{13, 14} Regarding specifically the use of equipment, an institution is obliged to assure that its employees are competent in the use of the equipment. Medical specialists have their own responsibility in maintaining demonstrable competence.¹⁷ This study shows that in practice this results in separated training of OR nurses and medical specialists as in almost all hospitals the equipment training was aimed at the OR nurses. Residents and medical specialists were mostly entitled as a separate group which organized their own training sessions. Team training sessions with the use of simulated scenarios can have benefits over separated training sessions as it gives the participants insight in the tasks of the different members of a team. This form of team training is applied in anesthesiology for some time with success.¹⁸ In one third of the hospitals training was not mandatory and in only a small number of hospitals assessment was included with the training sessions. Often, attending a training session is considered to be a warrant of competence, although evidently attendance on itself is not sufficient. In the current time of competency-based residency programs and stricter demands for the quality of healthcare, issues as described above deserve more attention.

CONCLUSION

Dutch hospitals pay attention to safety regarding laparoscopic operating equipment by the performance of preoperative equipment checks and training with the equipment. More than half of all hospitals use a checklist for preoperative checking of the equipment. The level of detail of these checklists varies and is often insufficient to prevent the occurrence of incidents with the equipment. In almost all hospitals training with equipment takes place, but a strict organization of these trainings is often lacking.

Broader implementation of detailed preoperative checklists for operating equipment, mandatory training and subsequent assessment of competence are preferable developments to assure safe en responsible use of operating equipment.

REFERENCES

- Risico's minimaal invasieve chirurgie onderschat 2007. Available from: http://www.igz. nl/15451/475693/2007-11_Rapport_Minimaal_in1.pdf.
- 2. Verdaasdonk EG, Stassen LP, van der Elst M, Karsten TM, Dankelman J. Problems with technical equipment during laparoscopic surgery. An observational study. Surg Endosc. 2007;21(2):275-9.
- 3. Courdier S, Garbin O, Hummel M, Thoma V, Ball E, Favre R, et al. Equipment failure: causes and consequences in endoscopic gynecologic surgery. J Minim Invasive Gynecol. 2009;16(1):28-33.
- 4. Persoon MC, Broos HJ, Witjes JA, Hendrikx AJ, Scherpbier AJ. The effect of distractions in the operating room during endourological procedures. Surg Endosc. 2011;25(2):437-43.
- Langelaan M, Baines RJ, Broekens MA, Siemerink KM, van de Steeg L, Asscheman H, et al. Monitor zorggerelateerde schade 2008: Dossieronderzoek in Nederlandse ziekenhuizen. Utrecht, The Netherlands. 2010. p. 1-120 http://www.nivel.nl/pdf/Rapport-zorggerelateerde-schade.pdf.
- 6. Inspectie voor de Gezondheidszorg (IGZ). Risico's van medische technologie onderschat. Den Haag: Inspectie voor de Gezondheidszorg (IGZ). 2008. p. 1-81.
- 7. Inspectie voor de Gezondheidszorg (IGZ). Standaardisatie onmisbaar voor risicovermindering in operatief proces. Den Haag: Inspectie voor de Gezondheidszorg (IGZ). 2008. p. 1-80.
- 8. Hales BM, Pronovost PJ. The checklist--a tool for error management and performance improvement. J Crit Care. 2006;21(3):231-5.
- 9. Hales B, Terblanche M, Fowler R, Sibbald W. Development of medical checklists for improved quality of patient care. Int J Qual Health Care. 2008;20(1):22-30.
- 10. Civil Aviation Authority (CAA). Guidance on the Design, Presentation and Use of Electronic Checklists. CAP 708. 2005 www.caa.co.uk/docs/33/CAP708.PDF.
- 11. Verdaasdonk EG, Stassen LP, Widhiasmara PP, Dankelman J. Requirements for the design and implementation of checklists for surgical processes. Surg Endosc. 2009;23(4):715-26.
- 12. Reason J. Understanding adverse events: human factors. Qual Health Care. 1995;4(2):80-9.
- 13. Nederlandse Vereniging voor Anesthesiologie (NVA), Nederlandse Vereniging voor Heelkunde (NVvH). Richtlijn Het Preoperatieve Traject. Utrecht. 2010. p. 1-63.
- 14. Nederlandse Vereniging voor Anesthesiologie (NVA), Nederlandse Vereniging voor Heelkunde (NVvH). Richtlijn Het Peroperatieve Traject. Utrecht. 2011. p. 1-37.
- 15. Meijer DW. Safety of the laparoscopy setup. Minim Invasive Ther Allied Technol. 2003;12(3):125-8.
- 16. Verdaasdonk EG, Stassen LP, Hoffmann WF, van der Elst M, Dankelman J. Can a structured checklist prevent problems with laparoscopic equipment? Surg Endosc. 2008;22(10):2238-43.
- 17. Orde van Medisch Specialisten. Verantwoordelijkheid medisch specialist bij onderhoud en beheer van medische apparatuur. Leidraad. Utrecht. 2008.
- 18. Waldrop WB, Murray DJ, Boulet JR, Kras JF. Management of anesthesia equipment failure: a simulation-based resident skill assessment. Anesth Analg. 2009;109(2):426-33.

CHAPTER 4.2

Effect of basic laparoscopic skills courses on essential knowledge of equipment

P.D. van Hove^{1,2}, E.G.G. Verdaasdonk³, E. van der Harst⁴, F.W. Jansen⁵, J. Dankelman¹, L.P.S. Stassen⁶

1: Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands 2: Department of Surgery, Reinier de Graaf Group, Delft, the Netherlands 3: Department of Surgery, Erasmus Medical Center, Rotterdam, the Netherlands 4: Department of Surgery, Maasstad Ziekenhuis, Rotterdam, the Netherlands 5: Department of Gynaecology, Leiden University Medical Center, Leiden, the Netherlands 6: Department of Surgery, Maastricht University Medical Center, Maastricht, the Netherlands

Published in: Surg Innov. 2012 Dec;19(4):460-6

ABSTRACT

Background Incidents with technical equipment during laparoscopy are in part caused by a lack of knowledge. This study aims to evaluate the effect of laparoscopic skills courses on the knowledge of laparoscopic equipment.

Methods A knowledge test on laparoscopic equipment was developed and participants of three separate basic laparoscopic skills courses in The Netherlands completed the test at the beginning and end of these courses. All lectures and demonstrations during the courses were recorded on video to assess the matching of its contents with the items in the test. As a reference, the test was also completed by a group of laparoscopic experts by email.

Results In total 36 participants (64.3%) completed both the pre- and posttest. Overall the mean test score improved from 60.4% of the maximum possible score for the pretest to 68.4% for the posttest (difference of 8.0%, 95% Cl 5.0 - 10.9%, p<0.001). There were no significant differences in test scores between the three separate courses. However, the actual content varied among the courses. The correspondence of the test items with the course content varied from 47% to 69%. Although 30% of the participants had already received training for laparoscopic equipment in their own hospital, 92.5% wanted to receive more training. Twenty-eight experts completed the test with a mean score of 75.7%, which was significantly better than the posttest score of the course participants (difference of 7.3%, 95% Cl 1.8-12.8%, p=0.011).

Conclusions The laparoscopic skills courses evaluated in this study showed to have a modest positive effect on the acquisition of knowledge of laparoscopic equipment. Variance exists among their contents. Further optimization of course content is important considering the crucial role of these courses in specialist training programs and the burden on financial resources.

INTRODUCTION

Presently, increasing amounts of technical equipment are being used during laparoscopic surgery. Unfortunately, this tendency is accompanied by preventable errors^{1, 2}. Verdaasdonk, in 2007, was the first to quantify the amount of incidents with laparoscopic equipment and reported one or more incidents during 26 of 30 laparoscopic procedures (87%)³. In 2008, Courdier published a similar study on a larger series of 116 endoscopic interventions in which he found one or more problems with the equipment during 45 interventions (39%)⁴. Both authors found that only a minor part of the malfunctions was caused by a defect and that human error contributed in most cases.

One of the contributing factors to incidents is inadequate knowledge of proper equipment handling, but also a lack of understanding the technology. For instance, several studies have indicated that specialists' and residents' knowledge of electrosurgical equipment is insufficient⁵⁻⁸. It is essential that a surgeon has sufficient knowledge of the equipment he works with, as he is often consulted by an OR nurse in case of a problem, but moreover, the surgeon is legally responsible for the result of the operation and the potential complications.

In The Netherlands, surgical residents are obligated to follow a basic laparoscopic skills course in the first year of their traineeship, before they are permitted to practice laparoscopic surgery as a primary surgeon. Most courses take one or two days outside the hospital and cost between 500 and 1000 euro's a person. It is known that basic laparoscopic courses have an improving effect on motor skills⁹, however there is no substantial evidence that this effect is also achieved for equipment handling in particular. Ideally, these courses should play an essential role in the improvement of knowledge and skills in order to practice laparoscopic surgery safely. A study by Condous showed improvement on knowledge of electrosurgery and the Harmonic scalpel during a hands-on course, but no other essential equipment was evaluated¹⁰. Moreover, Schaafsma showed that the more technically demanding (motor) skills, such as suturing, often distract the attention from other important basic skills during courses, of which equipment handling can be one¹¹.

Insight in what is the actual learning effect of laparoscopic skills courses on equipment handling is important in order to improve future training courses. The purpose of this study was to assess the effect of such courses on the acquisition of knowledge of technical surgical equipment.

METHODS AND MATERIALS

Test development

For this study a theoretical test was designed to assess knowledge of technical laparoscopic equipment. The test was designed by the primary researcher with the support of 4 general surgeons and one gynaecologist. These 5 specialists all perform advanced laparoscopic procedures regularly and are also experienced program and course directors. The questions in the test were made with the help of literature about equipment, user manuals and educational literature and websites¹²⁻¹⁹. The specialists approved the final version of the test.

The test considers the insufflator, light source, camera-unit, endoscope and electrosurgical equipment. It consists of 17 multiple choice questions, 12 statements which have to be answered with "true" or "false", and 3 multiple answer questions in which multiple correct answers should be selected. The questions consider the operating principles of the equipment, but also address physiological principles and consequences of insufflation and electrosurgery, possible complications from incorrect use of equipment and trouble shooting scenarios. All questions and statements of the test are shown in Tables 1 and 2.

Participants and scoring

The test was taken by all participants of 3 separate basic laparoscopic skills courses in The Netherlands at the start of the course (pretest) and at the end of the course (posttest). All 3 courses (A, B and C) were basic laparoscopic skills courses for residents at the beginning of their specialist training and with no or minimal laparoscopic experience. Course A and B were aimed at general surgical trainees and course C at gynaecological residents. The scores for the tests were calculated as a percentage of the total possible score and corrected for chance of guessing. For each question, the participants could score one point with a correct answer. For questions with multiple choice answers, the chance of guessing the right answer is 1 divided by the amount of answer possibilities, so for example ¼ for a question with 4 choices. Therefore, to correct for this chance, the sum of guessing chances for all questions was deducted from the total score and divided by the maximum possible score minus the sum of guessing chances. The result is the percentage of correctly answered questions, corrected for chance of guessing.

Course contents

The content of the courses was also evaluated. For this, all lectures and demonstrations given during the courses were recorded on video and scored at a later time for matching with the items in the test. The same was done for the self-study materials which the participants received in advance of the courses.

Table 1Multiple choice questions in the test.

For questions 18, 19 and 20 multiple correct answers had to be selected.

| | Questions |
|----|--|
| 1 | The most complete description of the function of an insufflator is: |
| 2 | With a pneumoperitoneum the intra abdominal pressure is increased. What is the effect of this increase in pressure on the cardiac preload? |
| 3 | What is a safe range for the intra abdominal pressure during a pneumoperitoneum at which you can expect a good exposure and minimal hemodynamic effects for a healthy adult with normal BMI? |
| 4 | What measures should an anesthaesiologist take to counter balance the physiological changes due to an increased intra abdominal pressure with a CO2 pneumoperitoneum? |
| 5 | What is the route of light and image with laparoscopy to produce the eventual image? |
| 6 | Which element in the chain for light and imaging is most important for the quality of the final image? |
| 7 | What is the best for the maintaining good quality of an image? Optical or digital zoom? |
| 8 | On insertion of the laparoscope it occurs to you that the colors have an unnatural appearance. Which of the following steps do you execute the first in order to improve the quality of the image? |
| 9 | Voltage can be described as: |
| 10 | Power can be defined as: |
| 11 | The essential difference between monopolar and bipolar electrosurgery is: |
| 12 | Different types of tissue offer different amounts of resistance to current. Which type of tissue offers the most resistance? |
| 13 | What is the consequence of the higher resistance of this tissue during electrosurgery? |
| 14 | With direct coupling: |
| 15 | With capacitive coupling: |
| 16 | The alarm for the neutral electrode of an electrosurgical device (Return Electrode Monitoring: REM) will go off when: |
| 17 | What is the most frequent occurring complication of electrosurgery in laparoscopic surgery? |
| 18 | Which parameters does a standard insufflator measure? |
| 19 | Select possible causes for activation of the alarm for excessive pressure. |
| 20 | Possible complications of long lasting excessive pressure in a pneumoperitoneum are: |

Questionnaire

At the end of the posttest all participants were asked to fill out a short supplementary questionnaire considering prior training with equipment and their opinion about knowledge of technical laparoscopic equipment.

Reference for test scores

As a reference for the scores of the course participants, experienced laparoscopic surgeons, working in hospitals throughout the Netherlands, were invited to take the same test. These experts had all performed more than 100 laparoscopic procedures. To assess their knowledge, the test was published as a questionnaire on a secured website and the experts were invited by an email with a link to the website. During one month, 2 remind-

| | Statements |
|----|---|
| 1 | The basic principle for intra abdominal pressure during a pneumoperitoneum is: the lower the pressure, the lower the risk of hemodynamic complications. |
| 2 | As the light used for laparoscopy is so-called 'cold light', it is impossible that the end of the laparoscope becomes hot. |
| 3 | When the image on the monitor is dark, it is NOT sufficient to adapt the clarity settings of the monitor. |
| 4 | The resolution is 'the amount off available pixels'. This is important for the quality of the image. |
| 5 | Voltages used with electrosurgery can increase to 5000 Volts. |
| 6 | With monopolar electrosurgery a lower voltage is usually applied than with bipolar electrosurgery. |
| 7 | For coagulation a lower voltage is usually used than for cutting. |
| 8 | For an optimal result with 'fulguration' the electrode must be in contact with the tissue. |
| 9 | If nothing happens when you step on the pedal for electrosurgery, the device is probably not producing enough power and therefore the first step should be to increase the power. |
| 10 | The use of metal trocars with a plastic anchor (hybrid trocars) prevents the occurring of capacitive coupling. |
| 11 | When operating on a patient with a pacemaker, it is safer to use bipolar electrosurgery instead of monopolar electrosurgery. |
| 12 | When operating on a patient with a metal implant, it is safest to place the neutral electrode close to this implant. |

Table 2Statements in the test to be answered with 'true' or 'false'.

ers were sent in order to get a good response. The scores for the tests were calculated in the same way as for the course participants.

Statistical Analysis

Data were analysed using SPSS 16.0 for Windows. For the course participants, the paired samples t-test was used to compare the differences between pre- and posttest scores and the ANOVA test with Bonferroni correction was used to test differences in posttest scores between the three different courses. To compare the scores of the course participants with the scores of the experts the independent samples t-test was used. A level of p<0.05 was considered to be statistically significant.

RESULTS

Knowledge of equipment

The total amount of participants for all 3 courses was 56, consisting of 11 participants for course A, 6 for course B and 39 for course C. Eventually, 36 participants (64%) fully completed both the pre- and posttest. Only the completed tests were used for analysis. General characteristics of the participants are presented in Table 3.

| | Overall | Course A | Course B | Course C |
|------------------------------------|--------------|--------------|--------------|--------------|
| Number of participants | 36 | 9 | 6 | 21 |
| Male : Female | 12:24 | 5:4 | 3:3 | 4:17 |
| Mean age (range) | 29 (26 - 33) | 28 (26 - 30) | 29 (27 - 31) | 30 (26 – 33) |
| Specialty of training | | | | |
| Gynaecology | 21 | 0 | 0 | 21 |
| General Surgery | 9 | 5 | 4 | 0 |
| Urology | 2 | 0 | 2 | 0 |
| Plastic Surgery | 2 | 2 | 0 | 0 |
| Orthopaedic Surgery | 1 | 1 | 0 | 0 |
| Thoracic Surgery | 1 | 1 | 0 | 0 |
| Year of residency | | | | |
| PGY 1 | 33 | 9 | 5 | 19 |
| PGY 2 | 3 | 0 | 1 | 2 |
| Performed laparoscopy as attending | | | | |
| surgeon before course | | | | |
| Yes | 13 | 1 | 5 | 7 |
| No | 23 | 8 | 1 | 14 |

 Table 3
 Characteristics of participants who completed both the pre- and posttest.

For the total group of participants (N=36) the mean score for the pretest was 60.4% of the maximum possible score. For the posttest the score increased to 68.4% (see Fig.1). The improvement of 8% is statistically significant (95% CI 5.0 - 10.9%, p<0.001). When separating results for the different courses (Fig.2 and Table 4), significant improvement was seen for every course (course A 10.0%, 95% CI 4.9 - 15.2%, course B 12.4%, 95%, CI 1.8 - 22.9%, course C 5.8%, 95% CI 1.8 - 9.9%). There were no significant differences between the posttest scores of the 3 courses.

The results were also analyzed for different parts of the test considering categories of equipment (insufflator, light & imaging and electrosurgery) (see Table 4). The posttest results did not differ much among the separate categories. For the pretest however, participants clearly scored the worst for electrosurgery. Consequently, this category showed a significant improvement of 12.9%. No significant improvement was found in the other two categories. A sub-analysis of all courses separately shows a similar pattern in every course (Table 4).

Expert group

In total, 29 of 36 (81%) invited experienced laparoscopic surgeons took the test. One participant was excluded from analysis because this participant turned out to have no experience as a laparoscopic surgeon. The results of the remaining 28 experts were analyzed. Characteristics of the experts are shown in Table 5.

The mean score for the expert group was 75.7%. This score was significantly higher than the posttest score of the course participants (difference of 7.3%, 95% Cl 1.8–12.8%,



Figure 1 Test results of pre- and posttests for all participants. Score is presented as a percentage of the maximum possible score.

Bars represent medians, boxes represent interquartile range and whiskers represent full range excluding outliers.

* Paired samples t-test.

p=0.011). The expert group scored 83.8%, 90.0% and 76.3% respectively for insufflator, light& imaging and electrosurgery.

Course contents

As presented in Table 6, the contents of the courses overlapped with the test items for 69, 50 and 47% respectively. Test items within the category 'light & imaging' were discussed the least frequent. For course A and B, most items were discussed during the course itself, but for course C most items were covered by the self-study materials.

Questionnaire

In total 40 participants (71.4%) fully completed the short questionnaire following the posttest. Of this group 92.5% indicated to want more training for equipment handling. The results of the questionnaire are presented in Table 7.





Bars represent medians, boxes represent interquartile range, whiskers represent full range excluding outliers and circles are outliers.

* Paired samples t-test.

DISCUSSION

To the authors' knowledge this is one of the first studies to focus in particular on the acquisition of knowledge of laparoscopic equipment during a skills course. The results of this study suggest a modest positive effect of basic laparoscopic skills courses on the acquisition of knowledge of technical laparoscopic equipment. Overall, the knowledge of participants of 3 courses increased with 8%. This overall significant improvement actually appears to be a reflection of a significant improvement in electrosurgical knowledge. However, scores for the insufflator and light & imaging parts were already fairly high in the pretest and were therefore harder to improve. The main question is what explains why only such a small improvement is achieved in overall knowledge of equipment. This could be caused either by insufficient discriminative capacity of the test to measure a difference, or there could be a deficit in the content of the courses.

To measure effectiveness of the courses purely for knowledge of technical laparoscopic equipment, a theoretical test was developed for this study. An official test specifically focused on the knowledge of laparoscopic equipment does not yet exist. The subject

| ** Corrected for chance | of guessing | E | | | | | | | | | | |
|-------------------------|-------------|----------|----------|---------|----------|----------|---------|----------|----------|---------|----------|----------|
| Equipment category | | Course / | A | | Course I | 8 | | Course (| | | Overall | |
| | pretest | posttest | p-value* | pretest | posttest | p-value* | pretest | posttest | p-value* | pretest | posttest | p-value* |
| Insufflator | 1.77 | 81.8 | 0.088 | 74.7 | 79.4 | 0.259 | 71.2 | 72.9 | 0.560 | 73.5 | 75.9 | 0.094 |
| Light & Imaging | 87.1 | 74.3 | 0.052 | 88.6 | 85.7 | 0.611 | 80.0 | 82.9 | 0.329 | 82.9 | 81.4 | 0.419 |
| Electrosurgery | 61.2 | 80.6 | 0.002 | 55.9 | 74.7 | 0.020 | 62.9 | 71.2 | 0.001 | 61.2 | 74.1 | <0.001 |
| Total | 72.4 | 80.0 | | 69.1 | 78.5 | | 69.3 | 73.8 | | 70.0 | 76.1 | |
| Corrected total** | 63.4 | 73.5 | 0.002 | 59.1 | 71.5 | 0.030 | 59.4 | 65.3 | 0.007 | 60.4 | 68.4 | <0.001 |
| | | | | | | | | | | | | |

Table 4 Scores for the pre- and posttests, given as a percentage of the maximum possible score. Values are mean percentages. Results are presented per category of equipment horizontally and are vertically presented separately for the pre and posttests. P-values are given for comparisons of scores for pre- and posttests. * Paired samples t-test.

| Characteristic | |
|--|--------------|
| Male : Female | 22:6 |
| Mean age (range) | 45 (34 – 60) |
| Specialism | |
| Surgery | 21 |
| Gynaecology | 2 |
| Urology | 5 |
| Mean amount of years as consultant (range) | 10 (1 – 26) |

| Table 5 | Characteristics of | expert group. |
|---------|--------------------|---------------|
|---------|--------------------|---------------|

Table 6Matching of course contents with the test items, given as percentages of test items discussedduring the three courses.

Categories of equipment are presented horizontally and self-study materials, the actual course and the combination of these two are presented vertically.

| Equipment category | C | ourse A | | c | ourse B | | C | ourse C | |
|-----------------------|------------|---------|---------|------------|---------|---------|------------|---------|---------|
| | self-study | course | overall | self-study | course | overall | self-study | course | overall |
| Insufflator | 0 | 50 | 50 | 0 | 50 | 50 | 63 | 13 | 63 |
| Light & Imaging | 0 | 43 | 43 | 0 | 29 | 29 | 14 | 0 | 14 |
| Electrosurgery | 53 | 65 | 88 | 59 | 35 | 59 | 47 | 18 | 53 |
| Total | 28 | 56 | 69 | 31 | 38 | 50 | 44 | 13 | 47 |

| Table 7 | Questions from the questionnal | re with the correspond | ling answers in | percentages. |
|---------|--------------------------------|------------------------|-----------------|--------------|
|---------|--------------------------------|------------------------|-----------------|--------------|

| Question | Answer | | |
|--|-------------------------|--------------------|---------------------------|
| | Yes (%) | No (%) | No opinion No idea (%) |
| Have you had any training for equipment handling in your hospital? If yes, did you consider that training sufficient? | 30.0 | 70.0 | 0.0 |
| | 41.7 | 58.3 | 0.0 |
| Would you like to have more training for equipment handling? | 92.5 | 7.5 | 0.0 |
| Do you use a checklist considering OR equipment? | 15.0 | 60.0 | 25.0 |
| Do you think a surgeon should have background knowledge about the operating room equipment he/she works with? | 100.0 | 0.0 | 0.0 |
| Do you think you should be able to answer all questions in the test correctly? | 82.5 | 15.0 | 2.5 |
| What did you think about the level of difficulty of the test? | Too difficult (%) | Good (%) | Too easy (%) |
| | 25.0 | 75.0 | 0.0 |

is addressed in part in the didactic modules and knowledge test of the Fundamentals of Laparoscopic Surgery (FLS) program of SAGES, but is not the main focus²⁰. To assure appropriateness of the content of the test, it was developed with the cooperation of 5 experienced laparoscopic surgeons. This was confirmed by the expert group scoring significantly better than the course participants, which is also proof of discriminative capacity of the test as experts are expected to have more knowledge than young residents. Additionally, 75% of the respondents indicated the level of difficulty to be 'good' and more than 80% indicated to think that they should be able to answer all questions of the test correctly (see Table 7).

The other explanation for the differences could be the course contents. Several deficits were shown in course contents. The percentage of discussed test items was equal to, or less than, 50% in 2 of the 3 evaluated courses. It is plausible that these deficits affect course effectiveness. Therefore, optimization of the course contents will presumably result in a larger effect on knowledge of technical laparoscopic equipment.

Another issue is the role of the teacher. At courses A and B an industrial representative was asked to give lectures about the equipment and also gave a full demonstration of the equipment at the start of the hands-on part of the courses. At course C all lectures were given by medical faculty and no demonstration of the equipment was given. Eventually in course A and B, more test items were discussed compared to course C (56% and 38% vs 13%). This could be the effect of a proper demonstration, or the expertise of the faculty.

The acquisition of knowledge which takes place during basic laparoscopic skills courses is of essential importance. The participants are at the start of their specialist training and are allowed to start practicing laparoscopy as a primary surgeon after completion of the course. Knowledge and understanding of equipment is equally important for safe surgery as surgical motor skills and decision making and should be addressed. Furthermore, the organization of these courses usually costs great efforts and financial support from multiple parties. For these reasons it is important that the quality of these courses is evaluated objectively to enable optimization. Therefore, assessment of what is learnt by participants is the key to improvement of courses.

CONCLUSION

The laparoscopic skills courses evaluated in this study, showed to have a moderate positive effect on knowledge of technical laparoscopic equipment. Moreover, variance exists among their contents. Further evaluation and optimization of these courses is of importance considering their crucial role in specialist training programs and the amounts of effort and money that is put into them.

REFERENCES

- 1. Brennan, T.A., L.L. Leape, N.M. Laird, et al., Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med. 1991;324:370-6.
- Leape, L.L., T.A. Brennan, N. Laird, et al., The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. N Engl J Med. 1991;324:377-84.
- 3. Verdaasdonk, E.G., L.P. Stassen, M. van der Elst, et al., Problems with technical equipment during laparoscopic surgery. An observational study. Surg Endosc. 2007;21:275-9.
- 4. Courdier, S., O. Garbin, M. Hummel, et al., Equipment failure: causes and consequences in endoscopic gynecologic surgery. J Minim Invasive Gynecol. 2009;16:28-33.
- 5. Mayooran, Z., S. Pearce, J. Tsaltas, et al., Ignorance of electrosurgery among obstetricians and gynaecologists. BJOG. 2004;111:1413-8.
- 6. Pandey, U., S.R. Killick, and S.W. Lindow, Obstetrics and gynaecology trainees' knowledge of electrosurgical equipment. J Obstet Gynaecol. 2007;27:721-2.
- 7. Sudhindra, T.V., A. Joseph, C.J. Hacking, et al., Are surgeons aware of the dangers of diathermy? Ann R Coll Surg Engl. 2000;82:31-2.
- 8. Menezes, C.A., D.W. Birch, A. Vizhul, et al., A deficiency in knowledge of basic principles of laparoscopy among attendees of an advanced laparoscopic surgery course. J Surg Educ. 2011;68:3-5 e2.
- 9. Schijven, M.P., J.J. Jakimowicz, I.A. Broeders, et al., The Eindhoven laparoscopic cholecystectomy training course--improving operating room performance using virtual reality training: results from the first E.A.E.S. accredited virtual reality trainings curriculum. Surg Endosc. 2005;19:1220-6.
- 10. Condous, G., D. Alhamdan, T. Bignardi, et al., The value of laparoscopic skills courses. Aust N Z J Obstet Gynaecol. 2009;49:312-5.
- 11. Schaafsma, B.E., E. Hiemstra, J. Dankelman, et al., Feedback in laparoscopic skills acquisition: an observational study during a basic skills training course. Gynecol Surg. 2009;6:339-343.
- 12. Broeders, I.A.M.J. and S. Kalisingh, eds. Handboek Endoscopische Chirurgie. 2009, Bohn Stafleu van Loghum: Houten.
- 13. Van Way, C.W., 3rd, Electrosurgery 101. Curr Surg. 2000;57:172-7.
- 14. Van Way, C.W. and C.S. Hinrichs, Electrosurgery 201: basic electrical principles. Curr Surg. 2000;57:261-264.
- 15. Meijer, D.W., Safety of the laparoscopy setup. Minim Invasive Ther Allied Technol. 2003;12:125-8.
- 16. Wu, M.P., C.S. Ou, S.L. Chen, et al., Complications and recommended practices for electrosurgery in laparoscopy. Am J Surg. 2000;179:67-73.
- 17. Marescaux, J., D. Mutter, N. Tanigawa, et al. World Electronic Book of Surgery WebSurg. Available at: www.websurg.com.
- 18. Valleylab and Covidien Electrosurgery Continuing Education Module. Available at: http://www. valleylabeducation.org/esself-2a/pages/esself2-01.html.
- 19. Medicines and Healthcare products Regulatory Agency MHRA Electrosurgery E-learning Module of the MHRA. Available at: http://www.mhra.gov.uk/ConferencesLearningCentre/LearningCentre/LearningCentre/Deviceslearningmodules/Electrosurgery/index.htm.
- 20. Peters, J.H., G.M. Fried, L.L. Swanstrom, et al., Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. Surgery. 2004;135:21-7.

CHAPTER 4.3

Validation of an interactive simulation module to train the use of a laparoscopic insufflator

P.D. van Hove^{1,2}, E.G.G. Verdaasdonk³, J. Dankelman¹, L.P.S. Stassen⁴

 Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands
 Department of Surgery, Maasstad Hospital, Rotterdam, the Netherlands
 Department of Surgery, Reinier de Graaf Group, Delft, the Netherlands
 Department of Surgery, Maastricht University Medical Center, Maastricht, the Netherlands

Accepted for publication in: Surgical Innovation

ABSTRACT

Background The purpose of this study was to determine construct and face validity of an interactive web-based module for pneumoperitoneum and insufflator.

Methods Participants were recruited from surgical departments in 2 academic hospitals and 1 large non-academic teaching hospital. They were stratified into 3 groups based on their laparoscopic experience (A: no experience, B: experience as assistant and C: experience as primary surgeon). Within each group the participants were randomized into a training subgroup and a control subgroup. All participants performed a theoretical and a practical test. The training participants first completed the module before they performed the tests. The control participants immediately performed the tests. Results were compared between the training and control participants. All training participants filled out a questionnaire on their opinion about the module.

Results In total 40 participants were enrolled in the study: group A consisted of 20 participants and group B and C both consisted of 10 participants. The trained participants answered significantly more theoretical questions correctly (8.3 vs 6.6; p<.001), correctly identified more alarm causes (91% vs 86%; p=.014) and made significantly less errors in the practical test (1.5 vs 3.6; p=.001). All 20 trained participants rated the module fairly good and indicated the module to be of additive value to surgical training programs.

Conclusions Training with the interactive web-based module on installation of a pneumoperitoneum and use of an insufflator has a positive effect on both theoretical and practical competence. Construct and face validity were established for this module.

INTRODUCTION

The increasing use of surgical technical equipment in the operating room has introduced new challenges for the medical personnel. Besides mastering operating skills, they also have to handle considerable amounts of equipment safely and prevent risks for patients and themselves. The need for such additional competencies has been illustrated by observational studies addressing incidents with operating equipment.¹⁻³ These incidents might induce risks and are a known source of delay and considerable distraction in the operating room.^{1, 2, 4}

Current education concerning equipment handling is often organized in a traditional way, using lectures and demonstrations given on site on an irregular basis.⁵ Such methods require considerable efforts in organization to ensure that all personnel is enabled to attend a training despite rotational schemes. However, attending such training sessions is often not mandatory. Moreover, assessment as a warrant of competency is often lacking.⁵ Outside the hospital, skills courses exist for residents and consultants in which equipment handling is expected to be part of the content. However, indications exist that such courses lack sufficient consideration of this subject.^{6,7}

In order to offer an easy-accessible training alternative for equipment handling, which can be used independently and includes assessment, an interactive web-based module was developed. The purpose of this study was to determine whether this module has construct and face validity.

MATERIALS AND METHODS

Development of the module

First a framework for the module and its training content was designed by two of the authors (PDVH and LS). The framework and content were discussed with various experts in the laparoscopic field (an anesthesiologist, a professor in laparoscopic surgery, a professor in pediatric surgery and a laparoscopic surgeon specialized in bariatric surgery). These experts all have wide experience in providing surgical training and laparoscopic courses. Then the actual web-based module was built with the help of a computer programmer and an interface designer from a company experienced in simulation for endoscopic surgical techniques (Simendo BV, Rotterdam, The Netherlands). A medical device production company (Olympus Netherlands BV, Zoeterwoude, Netherlands) agreed on the use of their insufflator as a model for a virtual simulated insufflator within the module.

The module consists of 3 parts: a theoretical part, a tutorial and an assessment. The theoretical part contains information about the physiological influences of a pneumo-

peritoneum on a human body, possible complications and the operating principles of an insufflator and its alarms. The tutorial consists of a practical explanation of all connections, buttons and indicators on an insufflator. An important component of the tutorial is a virtual simulation of a real insufflator (Olympus UHI-3), in which all connections and buttons are featured and a patient can be insufflated (Fig.1). This simulated device is also used in the assessment, in which the user encounters a number of scenarios on the simulator in which causes of alarms have to be identified and solved. These scenarios are alternated by theoretical (multiple choice) questions. At the end of the assessment the user receives a total score and sub-scores for 'theory pneumoperitoneum', 'theory insufflator', 'equipment handling' and 'trouble shooting'. Scores are given as a percentage of the maximum possible score.

Construct validity

The primary purpose of this study was to determine the construct validity of the module. The secondary purpose was to determine a possible effect of the experience of the participants on the effect of the module. Therefore a randomized controlled study was set up. All members of the surgical departments in 2 Dutch academic hospitals and one large non-academic teaching hospital were invited to participate in the study voluntarily. Upon participation an informed consent form was signed by each participant. Figure 2 shows a flowchart of the study design. The participants were first allocated to 1 of 3 different groups, based on their laparoscopic experience. Participants with no experience as an assisting or primary surgeon for laparoscopic procedures were allocated to group A, participants who only had experience as an assisting surgeon were allocated to group B and participants with experience as a primary surgeon during laparoscopic procedures were allocated to group C. After allocation to 1 of the 3 groups, the participants within each group were randomized into a training subgroup or a control subgroup

The participants in the training subgroup completed all three parts of the training module and after that performed a practical test on a real insufflator (Olympus UHI-3) to assess their practical competence. Their theoretical competence was extracted from the answers to the theoretical questions in the assessment within the module. The participants in the control subgroup did not have any training but started immediately with the theoretical test followed by the practical test on the real insufflator. The theoretical test for control participants was paper-based, but contained exactly the same questions as the assessment within the module. For the practical test the participants were given the assignment to: 1. establish all connections and actions to prepare the insufflator for use, 2. set the pressure at a given value, 3. set the flow rate at a given value and 4. start and stop the insufflation. Every assignment was divided into the required actions and each action was scored as either 'omission', 'completed correct' or 'completed incorrect' as well as the number of attempts. The total time needed for the whole practical test was measured as well. The practical test was exactly the



Figure 1 Two screenshots of the virtual simulation within the module, showing the trocars and the patient at the top and the front panel of the insufflator at the bottom.

same for training and control participants, as well as the content of the theoretical test, so participants in both groups performed exactly the same tests. To determine the construct validity of the module, the results on the theoretical and practical tests were compared between participants in the training subgroups and in the control subgroups. This was done for all participants of group A, B and C together and for the 3 groups separately.

Face validity

To evaluate face validity for the module, all participants in the training subgroups filled out a questionnaire after they had completed the module and the practical test. The questionnaire consisted of 27 questions addressing user-friendliness, level of difficulty, training capacity, level of realism and target group for the module.

Statistical analysis

All data were analyzed using the statistical software package SPSS for Windows, version 16.0 (SPSS Inc., Chicago, IL). A sample size was calculated based on the results of two pilot studies, which showed a mean difference of 8% (SD 11.5%) for the theoretical test and a mean of 0.78 (SD 0.86) errors for the practical test. Using these numbers in combination with a power of 80% and an α of 0.05, a sample size of 20 participants in both the training subgroup and the control subgroup was desired. For analysis of the training effect, the scores between training and control participants were compared using an independent samples t-test. A p-value smaller than 0.05 was considered statistically significant.

RESULTS

In total 40 participants were voluntarily enrolled in the study: 20 participants in group A and 10 participants in both group B and C. Within group A, B and C, the participants were





* The test for both groups consisted of a theoretical test and a practical test.

| | Gro | oup A | Gro | oup B | Gre | oup C | Ov | erall |
|--------------------|------------------|-----------------|-----------------|----------------|-----------------|----------------|------------------|-----------------|
| | Training n=10 | Control n=10 | Training n=5 | Control n=5 | Training n=5 | Control n=5 | Training n=20 | Control n=20 |
| Mean age | 24.2 | 25.2 | 26.4 | 25.8 | 33.6 | 37.8 | 27.1 | 28.5 |
| Gender (m:f) | 4:6 | 5:5 | 4:1 | 4:1 | 2:3 | 3:2 | 10:10 | 12:8 |
| Specialism | | | | | | | | |
| Surgery | 10 | 10 | 5 | 5 | 5 | 4 | 20 | 19 |
| Urology | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| BLS completed | 0 | 0 | 0 | 1 | 5 | 5 | 5 | 6 |
| Experience with in | sufflator as | assisting | urgeon | | | | | |
| 0 procedures | 10 | 10 | 0 | 0 | 0 | 0 | 10 | 10 |
| 1 – 20 procedures | 0 | 0 | 5 | 5 | 1 | 0 | 6 | 5 |
| >20 procedures | 0 | 0 | 0 | 0 | 4 | 5 | 4 | 5 |
| Experience with in | sufflator as | primary s | urgeon | | | | | |
| 0 procedures | 10 | 10 | 5 | 5 | 0 | 0 | 15 | 15 |
| 1 – 20 procedures | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 0 |
| >20 procedures | 0 | 0 | 0 | 0 | 3 | 5 | 3 | 5 |
| Worked with the sp | pecific insu | fflator used | d in the mo | dule in las | t two years | | | |
| Yes | 0 | 0 | 1 | 0 | 2 | 4 | 3 | 4 |
| No | 7 | 8 | 2 | 1 | 0 | 1 | 9 | 10 |
| Don't know | 3 | 2 | 2 | 4 | 3 | 0 | 8 | 6 |

Table 1 Characteristics of participants in the different groups.

BLS: Basic laparoscopic skills course

equally randomized into the training and the control subgroups. The characteristics of the participants are shown in Table 1.

Construct validity

The participants in the training subgroups scored better than those in the control subgroups for both the theoretical test and the practical test (Table 2). There were statistically significant differences for the mean amount of theoretical questions answered correctly and the mean amount of correctly identified alarm causes in the theoretical test and also for the mean amount of errors made during the practical test. The errors in the practical test largely consisted of repetitions of the same actions.

For the theoretical test the differences between the training and control subgroups were present over all groups in total and also in group A separately. In group B there was only a significant difference for the amount of correctly answered theoretical questions. In group C there were no differences between the training and the control subgroups.

For the practical test the differences in errors between training and control subgroups were present over all groups in total and in group A separately.

| Table 2 F | sesults for the | oretical and p | iractical test. | Numbers are | means ±SD. | | | | | | | |
|--|-------------------------------------|-------------------|-----------------|-----------------|----------------|----------|-----------------|----------------|----------|------------------|-----------------|----------|
| | | Group A | | | Group B | | | Group C | | | Overall | |
| | Training n=10 | Control n=10 | p-value* | Training n=5 | Control n=5 | p-value* | Training n=5 | Control n=5 | p-value* | Training n=20 | Control n=20 | p-value* |
| Theoretical | test | | | | | | | | | | | |
| Correctly answered theoretical questions (1 | 7.7 (±1.3) J | 5.9 (±1.1) | 0.003 | 8.9 (±0.2) | 7.3 (±1.2) | 0.04 | 8.7 (±0.5) | 7.3 (±1.5) | 0.082 | 8.3 (±1.1) | 6.6 (±1.3) | <0.001 |
| Correctly identified alarm cause (%)** | 90 (±4) si | 85 (±6) | 0.038 | 91 (±9) | 87 (±6) | 0.409 | 92 (土8) | 87 (±5) | 0.277 | 91 (±6) | 86 (±6) | 0.014 |
| Practical te: | tt | | | | | | | | | | | |
| Time (sec) | 152 (±59) | 170 (±32) | 0.412 | 139 (±31) | 148 (±61) | 0.781 | 161 (±88) | 169 (±91) | 0.896 | 151 (±60) | 164 (±55) | 0.480 |
| Errors | 1.6 (±1.5) | 4.0 (±2.1) | 0.009 | 1.2 (±1.3) | 3.6 (±3.0) | 0.136 | 1.6 (±1.7) | 2.8 (±1.3) | 0.242 | 1.5 (±1.4) | 3.6 (±2.1) | 0.001 |
| Repetitions | 1.5 (±1.6) | 3.1 (±1.8) | 0.048 | 1.2 (±1.3) | 2.8 (±2.6) | 0.252 | 1.6 (±1.7) | 2.6 (±1.1) | 0.302 | 1.5 (±1.5) | 2.9 (±1.8) | 0.008 |
| Incorrect execution | 0.0 (±0.0) | 0.1 (±0.3) | 0.343 | 0.0 (±0.0) | 0.4 (±0.5) | 0.178 | 0.0 (±0.0) | 0.2 (±0.4) | 0.374 | 0.0 (±0.0) | 0.2 (±0.4) | 0.042 |
| Omissions | 0.1 (±0.3) | 0.7 (±0.7) | 0.025 | 0.0 (±0.0) | 0.4 (±0.5) | 0.178 | 0.0 (±0.0) | 0.0 (土0.0) | 1.000 | 0.1 (±0.2) | 0.5 (±0.6) | 0.011 |
| * * | Inpaired t-test Corrected for fi | t alse answers | | | | | | | | | | |
| | | | | | | | | | | | | |

| Question | | Mean score (± SD) or N |
|--|----------------|------------------------|
| How do you rate the user-friendliness of the: | | |
| (1= very bad, 5= very good) | | |
| m | odule in total | 4.2 (± 0.7) |
| | simulator | 4.3 (± 0.6) |
| What do you think of the time needed to complete the: (1= too short, 3= exactly right, 5= too long) | | |
| | total module | 3.6 (± 0.8) |
| | theory part | 3.5 (± 0.7) |
| | tutorial | 3.2 (± 0.8) |
| | assessment | 3.4 (± 0.7) |
| How interesting or boring is the: (1= very boring, 5= very interesting) | | |
| | theory part | 4.3 (± 0.7) |
| | tutorial | 3.8 (± 0.8) |
| | assessment | 4.1 (± 0.8) |
| How do you rate the level of difficulty of the: (1= very easy, 5= very difficult) | | |
| | theory part | 3.2 (± 0.8) |
| | tutorial | 3.4 (± 0.8) |
| | assessment | 3.1 (± 0.6) |
| The assessment tests the content of the theory part and tutorial (1= not at all, 5= completely) | | 4.3 (± 0.7) |
| How do you rate the training capacity of the: (1= very bad, 5= very good) | | |
| | theory part | 4.1 (± 0.7) |
| | tutorial | 4.0 (± 0.6) |
| | assessment | 4.1 (± 0.6) |
| How do you rate the module's training capacity for: (1= very bad, 5= very good) | | |
| theoretical background concerning physiology | y and physics | 4.2 (± 0.7) |
| theoretical background concerning the | he insufflator | 4.2 (± 0.7) |
| operation of t | he insufflator | 4.0 (± 0.8) |
| trouble shooting for t | he insufflator | 4.4 (± 0.7) |
| How do you rate the realism of: (1= not realistic at all, 5= very realistic) | | |
| the looks of the sime | ulated device | 4.5 (± 0.5) |
| the sounds of the sime | ulated device | 4.4 (± 0.6) |
| connecting the sime | ulated device | 3.9 (± 0.8) |
| operating the sime | ulated device | 4.2 (± 0.7) |
| I considered any deficit in realism to be disturbing (1= very disturbing, 5= not disturbing at all) | | 4.2 (± 0.9) |
| | | |

 Table 3
 Results for face validity, presented as a mean score on a 5-point Likert scale (± SD).

| Question | N |
|---|----|
| I am now capable of operating the real insufflator in a safe and responsible way: | |
| I could already do this | 0 |
| I could already do this, but I have improved my skills | 5 |
| I was not capable of doing this, but now I am | 14 |
| I was not capable of doing this and I still am not | 1 |
| Don't know | 0 |
| This module is most appropriate for: | |
| OR nurses | 15 |
| Interns | 6 |
| Junior residents | 17 |
| Senior residents | 6 |
| Consultants | 3 |
| This module is of additive value to surgical training programs | |
| Yes | 20 |
| No | 0 |
| No opinion | 0 |

Table 4Results for face validity, presented as a number of respondents(N-total= 20).

Face validity

All 20 trained participants completed the questionnaire for face validity. Overall the participants rated the user-friendliness, level of difficulty, training capacity and level of realism with high scores (Table 3). Most participants indicated the module to be most appropriate for OR nurses and junior residents and all participants indicated the module to be of additive value for surgical training programs (Table 4).

DISCUSSION

An interactive web-based module was developed to train both theoretical and practical competency concerning installation of a pneumoperitoneum and use of an insufflator. A prominent component of this module is a virtual simulation of a real insufflator for which both construct and face validity were established. To the authors' knowledge computer simulation of equipment has not been described in literature before. Real-time simulation of anesthetic equipment has been described with good results.⁸⁻¹⁰

A positive training effect of the module was demonstrated by differences between participants in the training subgroups and the control subgroups for the results of the theoretical and practical tests. In this way construct validity was established. These differences were most evident within the least experienced group. It is most likely that the effect of training decreases with the increase of experience. Group A consisted of the least experienced participants, so any training is expected to show the clearest effect in this group. Therefore, it was a deliberate choice to include more (inexperienced) participants in this group to reduce the effect of confounding by experience on evaluation the primary purpose of this study, which was to evaluate the effect of the module.

An insufflator is a rather simple device to operate, so a minimum of experience with such a device will have a clear effect on practical competence. Theoretical knowledge about a pneumoperitoneum or an insufflator is probably not as directly influenced by experience. This may explain why no difference was found for the practical test results between training and control participants in groups B and C, but there was still a difference within group B for the results of the theoretical test. However, due to the smaller number of participants in groups B and C, other differences may have remained unnoticed.

All participants who worked with the module rated it positively in the questionnaire for face validity. Most participants answered that they found the module most suitable for junior residents and for OR nurses. In the development process the theoretical content was specifically aimed at residents and consultants so in its present form the module is probably too difficult for OR nurses. However, some adaptations can easily be made, which will result in a separate module especially suitable for this group.

The influence of technology in the operating room is increasing and it is becoming a heavy burden for OR-personnel to stay up-to-date with all equipment. Previous research showed that training of OR-personnel with equipment is not as structured as it could be and often lacks assessment of competency.⁵ Likewise content of laparoscopic skills courses for residents may be lacking sufficient focus on equipment.⁶ Moreover, the official 48-hour European workweek for residents demands effective and efficient means for training. An easy accessible module, which can be used independently and which both trains and assesses theoretical and practical competence, could facilitate hospitals in training their personnel despite rotational schemes, to certify them for equipment-competency and to register these certifications. For the same reasons it would be a valuable addition to current training programs and skills courses. Another purpose of use, especially of the virtual simulation of a device, is the opportunity to get insight into the human-machine interaction. This may render valuable information for manufacturers and can help solve frequent occurring issues, as was proven earlier in anesthesiology.⁹

CONCLUSION

This study demonstrated a positive effect on theoretical and practical competence after training with an interactive web-based module for installation of a pneumoperitoneum and use of an insufflator. The effect was most clear in inexperienced subjects. All participants who trained with the module rated it positively and indicated it to be a useful

addition to surgical training programs. Thus construct and face validity were established for this module.

REFERENCES

- 1. Verdaasdonk EG, Stassen LP, van der Elst M, Karsten TM, Dankelman J. Problems with technical equipment during laparoscopic surgery. An observational study. Surg Endosc. 2007;21(2):275-9.
- 2. Courdier S, Garbin O, Hummel M, Thoma V, Ball E, Favre R, et al. Equipment failure: causes and consequences in endoscopic gynecologic surgery. J Minim Invasive Gynecol. 2009;16(1):28-33.
- 3. Panesar SS, Salvilla SA, Patel B, Donaldson SL. Laparoscopic cholecystectomy: device-related errors revealed through a national database. Expert Rev Med Devices. 2011;8(5):555-60.
- 4. Persoon MC, Broos HJ, Witjes JA, Hendrikx AJ, Scherpbier AJ. The effect of distractions in the operating room during endourological procedures. Surg Endosc. 2011;25(2):437-43.
- 5. van Hove PD, Verdaasdonk EG, Dankelman J, Stassen LP. [Use of checklists for laparoscopic equipment in Dutch hospitals]. Ned Tijdschr Geneeskd. 2012;156(33):A4505.
- 6. van Hove PD, Verdaasdonk EG, van der Harst E, Jansen FW, Dankelman J, Stassen LP. Effect of Basic Laparoscopic Skills Courses on Essential Knowledge of Equipment. Surg Innov. 2012.
- Menezes CA, Birch DW, Vizhul A, Shi X, Sherman V, Karmali S. A deficiency in knowledge of basic principles of laparoscopy among attendees of an advanced laparoscopic surgery course. J Surg Educ. 2011;68(1):3-5 e2.
- 8. Berge J.A, Gramstad L, Jensen O. A training simulator for detecting equipment failure in the anaesthetic machine. Eur J Anaesthesiol. 1993;10(1):19-24.
- 9. Mudumbai SC, Fanning R, Howard SK, Davies MF, Gaba DM. Use of medical simulation to explore equipment failures and human-machine interactions in anesthesia machine pipeline supply crossover. Anesth Analg. 2010;110(5):1292-6.
- 10. Waldrop WB, Murray DJ, Boulet JR, Kras JF. Management of anesthesia equipment failure: a simulation-based resident skill assessment. Anesth Analg. 2009;109(2):426-33.

CHAPTER 4.4

Development and evaluation of an interactive simulation module to train the use of an electrosurgical device

P.D. van Hove^{1,2}, E.G.G. Verdaasdonk³, J. Dankelman¹, L.P.S. Stassen⁴

 Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands
 Department of Surgery, Maasstad Hospital, Rotterdam, the Netherlands
 Department of Surgery, Reinier de Graaf Group, Delft, the Netherlands
 Department of Surgery, Maastricht University Medical Center, Maastricht, the Netherlands

Submitted

ABSTRACT

Background The purpose of this study was to develop an interactive web-based training module for electrosurgery and use of an electrosurgical device, and to evaluate its training effect and face validity.

Methods The training module was developed by a multidisciplinary team and consists of a theoretical part, a device tutorial and an assessment. For evaluation participants were recruited at the surgical departments from a university hospital and a non-university teaching hospital and were divided into a training group and a control group. All participants performed the same theoretical and practical test. The training participants first completed the module before they performed the tests. The control participants immediately performed the tests. Results were compared between the training and control participants. To evaluate face validity, the training participants filled out a questionnaire on their opinion about the module.

Results In total 39 participants were enrolled in the study: 20 in the training group and 19 in the control group. The training group answered significantly more theoretical questions correctly (15.7 vs 9.7; p< .001) and made significantly less errors in the practical test (2.2 vs 5.6; p= .007). The participants in the training group rated the usefulness and characteristics of the module with high marks. All of them indicated the module to be of additive value to surgical training programs.

Conclusions Training with an interactive web-based module has a positive effect on both theoretical and practical competence regarding electrosurgery and use of an electrosurgical device. This module was rated positively by the participants and was indicated to be a useful addition to surgical training programs.

INTRODUCTION

The increasing use of surgical technical equipment in the operating room has introduced new challenges for medical personnel. Besides mastering operating skills, they also have to handle increasing amounts of equipment safely and prevent risks for patients and themselves. The need for such additional competencies has been illustrated by observational studies addressing incidents with laparoscopic operating equipment.¹⁻³ Electrosurgical devices are one of the multiple types of equipment and are already being used for more than a century in surgery.^{4, 5} Although these devices have become much safer than they were, laparoscopic surgery introduced a new form of application and with that new sorts of accidents.⁶ The literature has reported accidents occurring with electrosurgery in laparoscopic surgery varying from 2 to 5 per 1000 cases.^{6, 7} Small studies have shown deficiencies in knowledge on electrosurgery among residents and medical specialists.⁸⁻¹¹

Current education concerning equipment handling, including electrosurgery, is often organized in a traditional way, using lectures and demonstrations given on site on an irregular basis.¹² Attending such training sessions is often not mandatory. Moreover, assessment as a warrant of competency is frequently lacking.¹² Outside the hospital, skills courses exist for residents and consultants in which instructions for and training of equipment handling is expected to be part of the content. However, indications exist that such courses lack sufficient consideration of this subject and therefore do not have the intended effect on knowledge and skills.¹³⁻¹⁵

In order to offer an easy-accessible training alternative for equipment handling, which can be used independently and includes assessment, the purpose of this study was to develop an interactive web-based training module on electrosurgery and the use of an electrosurgical device. The second purpose of this study was to evaluate this module's training effect and face validity.

METHODS

Development of the module

First a framework for the module and its training content was designed by the first author. The framework and content were discussed with several experts among which one of the other authors (LPSS) and a field expert from a company that produces and sells electrosurgical devices (ERBE Netherlands, Werkendam, The Netherlands). Then the actual web-based module was built by a computer programmer and an interface designer from a company experienced in simulation for endoscopic surgical techniques (Simendo BV, Rotterdam, The Netherlands). The module consists of 3 parts: a theoretical part, a device tutorial and an assessment. The electrosurgical device producing company agreed on the use of one of their electrosurgical devices (ERBE VIO 300D) as a model for a virtual simulated device within the module.

Evaluation of training effect

To evaluate the training effect of the module under study, a prospective study was set up with a training group and a control group. Participants were recruited from the surgical departments of a university hospital and a non-university teaching hospital. Members of the surgical department of the non-university teaching hospital were invited to participate in the training group and members of the surgical department of university hospital were invited to participate in the training department of the control group. All participants volunteered and signed an informed consent upon participation.

The participants in the training group completed all 3 parts of the training module and after that performed a separate practical test on the real electrosurgical device (ERBE VIO 300D) to assess their practical competence. Their theoretical competence was extracted from answers to the theoretical questions in the assessment within the module. The participants in the control group did not have any previous training but started immediately with the theoretical and the practical test. The theoretical test for the control group was paper-based, but contained exactly the same questions as the assessment within the module. These questions have been validated in a previous study.¹³ For the practical test the participants were given the assignment to: 1. Connect the generator to a power source 2. establish all connections to prepare the generator for monopolar diathermy, 3. establish extra connections for bipolar diathermy 4. adjust several settings for both monopolar and bipolar diathermy for given scenarios. Every assignment was divided into the required actions and each action was scored as either 'completed correct', 'completed incorrect' or 'omission' as well as the number of attempts. The total time needed for the whole practical test was measured as well. The practical test was exactly the same for training and control participants. To determine the training effect of the module, the results for the theoretical and practical tests were compared between the training group and the control group.

Face validity

Face validity refers to a subjective evaluation of a training method's content and its resemblance to reality.^{16,17} To evaluate face validity for the module in this study, all participants in the training group filled out a questionnaire after they had completed the module and the practical test. The questionnaire consisted of 27 questions addressing user-friendliness, level of difficulty, training capacity, level of realism and target group for the module.

Statistical analysis

All data were analyzed using the statistical software package SPSS for Windows, version 16.0 (SPSS Inc., Chicago, IL). For analysis of the training effect, the scores between training and control participants were compared using an independent samples t-test. A p-value smaller than 0.05 was considered statistically significant. To verify the sample size a power analysis was performed. A difference in means of 3.4 errors for the practical test with a within group standard deviation of 2.1 and a target significance level of 0.05 for 19 training participants and 19 control participants achieves a power of 99.7%.

RESULTS

The module

The module consists of 3 parts: a theoretical part, a device tutorial and an assessment. The theoretical part contains information about the physical background of electrosurgery, the operating principles and the possible complications of the use of an electrosurgical generator and its alarms. The tutorial consists of a practical explanation of all connections, buttons, indicators and menus of the device including illustrative short videos. An important component of the tutorial is a virtual simulation of the real electrosurgical device in which all connections, buttons and electrodes are featured (Fig.1). In the tutorial this simulation can be used for free practice. In the assessment it is used for scenarios in which the user has to establish several connections and settings. These scenarios are alternated by theoretical (multiple choice) questions. At the end of the assessment the user receives a total score and sub-scores for 'theory electrosurgery', 'theory electrosurgical generator', 'equipment handling' and 'trouble shooting'. Scores are given as a percentage of the maximum possible score.

Evaluation

In total 39 participants were voluntarily enrolled in the study: 20 participants in the training group and 19 participants in the control group. Table 1 shows that there were no important differences between the two groups, except for the fact that the control group contained more participants who were experienced with an electrosurgical device from ERBE than the training group.



Figure 1 Screenshots of the simulated electrosurgical device within the module, showing the front of the device and an enlargement of its screen.

| | Gre | Group | |
|--|-----------------|----------------|--|
| | Training (N=20) | Control (N=19) | |
| Mean age | 27 | 28 | |
| Gender (m:f) | 12:8 | 11:8 | |
| BLS completed | 6 | 5 | |
| Experience as an attending surgeon | | | |
| 0 procedures | 13 | 11 | |
| 1 – 20 procedures | 1 | 2 | |
| >20 procedures | 6 | 6 | |
| Worked with the specific device used in the module in last 2 years | | | |
| Yes | 0 | 5 | |
| No | 11 | 7 | |
| Don't know | 9 | 7 | |

Table 1 Characteristics of participants in the different groups.

BLS: Basic laparoscopic skills course

Training effect

The participants in the training group outperformed those in the control group for both the theoretical test and the practical test (Table 2). There was a statistically significant difference for the mean amount of theoretical questions answered correctly in the theoretical test and for the mean amount of errors made during the practical test. The errors in the practical test mostly consisted of repetitions of the same actions.

| | Training group N = 20 | Control group N= 19 | p-value* |
|---|--------------------------|------------------------|----------|
| Theoretical test | | | |
| Correctly answered theoretical questions (N) ** | 15.7 (±1.8) | 9.7 (±2.4) | <0.001 |
| Practical test | | | |
| Time (seconds) | 185 (±43) | 220 (±70) | 0.063 |
| Errors | 2.2 (±2.1) | 5.6 (±4.7) | 0.007 |
| Repetitions | 2.1 (±2.1) | 5.2 (±4.2) | 0.006 |
| Incorrect execution | 0 | 0 | - |
| Omissions | 0.1 (±0.4) | 0.4 (±1.4) | 0.344 |

| Table 2 | Results for theoretical and | practical test. Numbers are means ±SD. |
|---------|-----------------------------|--|
|---------|-----------------------------|--|

* Unpaired t-test

** Maximum possible score was 18

Face validity

Of all 20 participants in the training group, 16 fully completed the questionnaire for face validity. One participant did not complete the questionnaire at all and 3 other participants did not fill out part of the questions. Overall the participants rated the user-friendliness, level of difficulty, training capacity and especially level of realism with high scores (Table 3). Most participants indicated the module to be most appropriate for OR nurses and junior residents and all participants indicated the module to be of additive value for surgical training programs (Table 4).

DISCUSSION

This paper discusses the development and evaluation of a newly designed web-based module to train both theoretical and practical competence regarding electrosurgery and handling electrosurgical equipment. The evaluation of the module demonstrated a positive training effect as the training group outperformed the control group significantly for both the theoretical and the practical test. This confirms the hypothesis that it is possible to train practical equipment-related competence without the need of an actual device. This offers new opportunities for training and assessment of healthcare personnel. As it has been shown that incidents with technical equipment are not infrequent¹⁻³ and as requirements by healthcare inspectorates are sharpening, there is a constant need to keep personnel's skills and knowledge up-to-date. Not only is there a need for adequate training, but healthcare institutions are required to be able to show proof of the competence of their personnel. Web-based modules, like the one described and tested in this paper, might offer a mean to live up to such requirements and to offer personnel an easy-to-use alternative for independent training and assessment without the need of an instructor, assessor or the

| Question | Mean score (± SD) or N |
|--|------------------------|
| How do you rate the user-friendliness of the: | |
| (1= very bad, 5= very good) | |
| module in total | 4.0 (±0.7) |
| simulator | 3.8 (±0.9) |
| What do you think of the time needed to complete the: (1= too short, 3= exactly right, 5= too long) | |
| total module | 3.3 (±0.8) |
| theory part | 3.4 (±0.7) |
| tutorial | 3.2 (±0.7) |
| assessment | 3.5 (±0.8) |
| How interesting or boring is the: (1= very boring, 5= very interesting) | |
| total module | 3.4 (±0.9) |
| theory part | 3.3 (±1.1) |
| tutorial | 3.7 (±0.8) |
| How do you rate the level of difficulty of the: (1= very easy, 5= very difficult) | |
| theory part | 3.1 (±0.8) |
| tutorial | 3.2 (±0.9) |
| assessment | 2.8 (±0.7) |
| The assessment tests the content of the theory part and tutorial (1= not at all, 5= completely) | 4.0 (±1.0) |
| How do you rate the training capacity of the: (1= very bad, 5= very good) | |
| total module | 3.9 (±0.7) |
| theory part | 3.7 (±0.7) |
| tutorial | 3.7 (±0.8) |
| How do you rate the module's training capacity for: (1= very bad, 5= very good) | |
| theoretical background concerning physiology and physics | 3.8 (±0.8) |
| theoretical background concerning the device | 3.8 (±1.0) |
| practical handling of the device | 4.0 (±0.8) |
| trouble shooting for the device | 3.4 (±0.9) |
| How do you rate the realism of: (1= not realistic at all, 5= very realistic) | |
| the looks of the simulated device | 4.4 (±0.8) |
| the sounds of the simulated device | 4.2 (±0.8) |
| connecting the simulated device | 4.5 (±0.8) |
| operating the simulated device | 4.3 (±0.8) |
| l considered any deficit in realism to be disturbing (1= very disturbing, 5= not disturbing at all) | 4.3 (±1.1) |

| Table 3 Results for face validity, presented as a mean | n score on a 5-point Likert scale (± SD). |
|--|---|
|--|---|
availability of equipment. Likewise, such modules can be applied as preparation for skills courses. In a previous study our group showed that laparoscopic skills courses only have a slightly positive effect on participants' knowledge of laparoscopic equipment.¹³ When course participants could use a web-based module for home study in advance of a skills course, the eventual effect of a course, both theoretically and practically, may increase.

Participants in the training group gave high scores for user-friendliness, training capacity and realism. Moreover, all participants who filled out the questionnaire indicated the module to be of additive value to surgical training programs. The participants indicated they thought the module is most suitable for OR nurses and junior residents. In its current form, the module was designed for residents and medical specialists and to the authors' opinions contains too much detailed theoretical information for OR nurses. However, adaptations in its content can be made reasonably easy and will make it more appropriate for other groups and thereby expand its applicability.

There is a limitation to this study in the fact that participants were not randomly divided into groups, but that the 2 study groups were recruited at two different hospitals. This is a potential source of bias. Indeed, there was a difference in familiarity with the specific device used in the module. In the control group more participants had worked with this device in the last 2 years (Table 1). Despite this difference, the training group still outperformed the control group, which actually underlines the positive training effect of the module on practical competence even more. There were no other differences in prior training related to electrosurgery which could have interfered with this effect.

According to the results of this study, the training module can be used to train theoretical and practical competence regarding electrosurgery. However, as this was only a first evaluation of the effect of this module, some questions remain unanswered. For instance, it is not yet clear for which exact groups the module works best. The assessment of the module was not validated in this study, so no conclusions can be drawn from its scores yet. Therefore, at this stage, the module should only be used for voluntary training with no consequences related to the results of its assessment. Future research should aim at further distinct evaluation of the exact target group and validation of the assessment together with determination of cut-off scores for 'pass' or 'fail'.

CONCLUSION

An interactive web-based training module was developed for both theoretical and practical competence regarding knowledge of electrosurgery and use of electrosurgical equipment. This study demonstrated a positive effect on theoretical and practical competence after training with the module. All participants who trained with the module rated it positively and indicated it to be a useful addition to surgical training programs.

REFERENCES

- 1. Verdaasdonk EG, Stassen LP, van der Elst M, Karsten TM, Dankelman J. Problems with technical equipment during laparoscopic surgery. An observational study. Surg Endosc. 2007;21(2):275-9.
- 2. Courdier S, Garbin O, Hummel M, Thoma V, Ball E, Favre R, et al. Equipment failure: causes and consequences in endoscopic gynecologic surgery. J Minim Invasive Gynecol. 2009;16(1):28-33.
- 3. Panesar SS, Salvilla SA, Patel B, Donaldson SL. Laparoscopic cholecystectomy: device-related errors revealed through a national database. Expert Rev Med Devices. 2011;8(5):555-60.
- 4. Beer E. Removal of neoplasms of the urinary bladder. A new method employing high frequency (Oudin) currents through a catheterizing cystoscope. JAMA. 1910;54(1768.
- 5. Cushing H. Electrosurgery as an aid to the removal of intracranial tumors. Surg Gynecol Obstet. 1928;47(751 84.
- 6. Nduka CC, Super PA, Monson JR, Darzi AW. Cause and prevention of electrosurgical injuries in laparoscopy. J Am Coll Surg. 1994;179(2):161-70.
- 7. Wu MP, Ou CS, Chen SL, Yen EY, Rowbotham R. Complications and recommended practices for electrosurgery in laparoscopy. Am J Surg. 2000;179(1):67-73.
- 8. Pandey U, Killick SR, Lindow SW. Obstetrics and gynaecology trainees' knowledge of electrosurgical equipment. J Obstet Gynaecol. 2007;27(7):721-2.
- 9. Assiotis A, Christofi T, Raptis DA, Engledow A, Imber C, Huang A. Diathermy training and usage trends among surgical trainees--will we get our fingers burnt? Surgeon. 2009;7(3):132-6.
- Sudhindra TV, Joseph A, Hacking CJ, Haray PN. Are surgeons aware of the dangers of diathermy? Ann R Coll Surg Engl. 2000;82(1):31-2.
- 11. Tucker RD. Laparoscopic electrosurgical injuries: survey results and their implications. Surg Laparosc Endosc. 1995;5(4):311-7.
- 12. van Hove PD, Verdaasdonk EG, Dankelman J, Stassen LP. [Use of checklists for laparoscopic equipment in Dutch hospitals]. Ned Tijdschr Geneeskd. 2012;156(33):A4505.
- 13. van Hove PD, Verdaasdonk EG, van der Harst E, Jansen FW, Dankelman J, Stassen LP. Effect of Basic Laparoscopic Skills Courses on Essential Knowledge of Equipment. Surg Innov. 2012.
- Menezes CA, Birch DW, Vizhul A, Shi X, Sherman V, Karmali S. A deficiency in knowledge of basic principles of laparoscopy among attendees of an advanced laparoscopic surgery course. J Surg Educ. 2011;68(1):3-5 e2.
- 15. Mayooran Z, Pearce S, Tsaltas J, Rombauts L, Brown TI, Lawrence AS, et al. Ignorance of electrosurgery among obstetricians and gynaecologists. BJOG. 2004;111(12):1413-8.
- 16. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. Surg Endosc. 2003;17(10):1525 9.
- 17. Schijven MP, Jakimowicz JJ. Validation of virtual reality simulators: Key to the successfull integration of a novel teaching technology into minimal access surgery. Minim Invasive Ther Allied Technol. 2005;14(4-5):244 - 6.

CHAPTER 5

General conclusions, discussion and future perspectives

GENERAL CONCLUSIONS

The objective of this thesis was to improve laparoscopic surgical skills training and assessment, with an emphasis on equipment-related safety and competence.

With regard to the objective assessment of surgical skills, in the first part of this thesis the literature was reviewed. This showed that there is currently no available method which is valid and reliable enough to be used for summative assessment of skills in the operating room. With regard to laparoscopic surgical skills training this thesis evaluated an 'intermediate skills' curriculum on a virtual reality simulator for laparoscopic skills. This curriculum for advanced skills training was shown to discriminate in level of performance of advanced laparoscopic skills. This suggests that such a VR curriculum might also be used to train these skills.

In the second part of this thesis a comparison was made between healthcare and the petrochemical industry on different aspects. Safety management systems and safety culture are in an immature stage in healthcare, when compared to the petrochemical industry. The comparison of a hospital and a petrochemical company showed that equipment-related training of personnel was managed very strictly in the petrochemical company compared to the studied hospital. In the hospital training with equipment was not mandatory, did not include assessment and competent personnel was not certified.

The third part of this thesis emphasized equipment-related safety and training. A questionnaire revealed that in 56% of the Dutch hospitals a preoperative checklist, specifically aimed at technical laparoscopic equipment, is used. Training with this equipment takes place in nearly all hospitals but varies highly in frequency. Moreover, training is not mandatory in one third of all hospitals in The Netherlands and does not include an assessment of competence. Basic laparoscopic skills courses for surgical residents were shown to improve essential knowledge of technical laparoscopic equipment by a moderate 8%. The contents of different courses appeared to be variable and cover only a part of essential topics about equipment. Finally, two interactive web-based training modules for laparoscopic equipment were developed and evaluated. These modules considered an insufflator and an electrosurgical generator. Both modules showed a significant effect of training on theoretical and practical competence and were well appreciated by the study participants.

GENERAL DISCUSSION AND FUTURE PERSPECTIVES

Technology is of major influence in healthcare and there is an ongoing development. In surgery this is accompanied especially with an urge to minimize the size and number

of incisions. Such minimally invasive surgery is propagated to have advantages over conventional techniques. However, this technique has one important drawback: if the surgical technique and handling of the accompanying equipment are not mastered, they pose a major threat to patient safety. This is the justification for the need of training methods, assessment of competence and systems for certification and registration of those who want to practice these techniques.

Assessment of laparoscopic competence

In its report on minimally invasive surgery in 2007, the Dutch Healthcare Inspectorate (DHI) suggested implementation of intercollegiate assessment of competence among laparoscopic surgeons by video recordings.¹ This should facilitate certification of laparoscopic surgeons. At present such a method is not yet in use, because a method valid and reliable enough for certification does not exist. In the development of a suitable method for assessment several complex issues have to be overcome.

The first issue is to determine what exactly is to be assessed. Generally, skills assessment can be divided into a general approach or a procedure-specific approach. A general approach is focused on general skills like tissue handling or use of instruments, but also parameters such as task time or motion analysis parameters are general measures of skills. With a procedure-specific (or procedure-based) approach the different steps of one specific procedure are defined and rated with 'good' or 'false'. Several national surgical training programs currently make use of the procedure-specific approach as a feedback tool² under the belief that this method gives the most reliable results.

With a suitable method at hand, a second issue is to determine what is to be scored as 'good' or 'false'. Surgical skills are means to reach a final result. However, there may be different ways to reach the same result. Certain approaches might not be 'the standard approach', but these are not necessarily wrong. Herein lays the risk of inter-rater variability of an assessment. A reliable assessment has a low inter-rater variability. One way to reduce inter-rater variability is the use of 'best practice protocols'.³ These are guidelines developed by surgical societies, which contain a consensus on the essential steps of an operation. Such documents could form a fundament for assessment tools for these specific procedures. Unfortunately, best practice protocols do not exist for every operation and are usually restricted to several important steps of an operation.

As certification can have major consequences, a 'pass' or 'fail' judgment of an assessment method needs to be undisputable. Therefore, a method should be able to depict all the essential information; i.e. the sensitivity of that method. This does not only depend on the method itself, but also on the form in which it is applied. For instance, a procedure-specific rating scale can have a high sensitivity when used live in the operating room, but might not reach this level of sensitivity when used in combination with video recordings as essential information might be missed on the video recordings.⁴ Consequently, the right method should be used in the right environment.

The issues discussed above might seem straight forward. However, to solve them in developing a valid and reliable tool for intercollegiate assessment of surgical skills is complex. Nonetheless, a research group from Imperial College London has recently succeeded in developing, validating and implementing a reliable and highly sensitive procedure-based assessment tool for colorectal surgery. This tool is incorporated in a national training program for colorectal surgery and is used for summative judgment of the competence of apprentices at the end of the training program.⁵ The results of this group indicate that assessment of specialist performance on a national level is eventually possible and that procedure-based assessment is the way to go. Future work in this area should focus on the development of procedure-based assessment tools for different procedures and the adoption of these tools in other countries and training programs.

Equipment-related competence

Whereas laparoscopic surgical skills training has received abundant attention in the past decade, training related to equipment has not. The fact that technical equipment interferes with safety is not denied in the many reports that have been written by authorities, but most of them mainly address the purchase and maintenance processes.⁶⁻⁸ However, the true man-machine interaction between the equipment and its user is actually most relevant in this discussion⁹. According to several documents, in The Netherlands, hospitals have to ascertain that their personnel are competent in using the equipment required for their job.¹⁰⁻¹² Nonetheless, as shown in this thesis (chapter 3.2 and 4.1), hospitals do not seem to be eager to fulfill this requirement. The primary reason for this is probably that training with equipment is not prioritized (during audits) by authorities. Conclusively, there is a need to improve training with equipment and there is definitely room to achieve this.

The first point of action is to make training with equipment an urgent topic. Therefore, hospital boards need to propagate training by making it mandatory and set up frameworks for regular training cycles and certification. This can be stimulated by authorities when the topic is prioritized more during audits and when medical societies address it more explicit in their guidelines. Furthermore, to ensure competency of personnel, not only improved training is required but a whole system of training, assessment, certification and registration. To set up such a systematic approach requires manpower. Chapter 3.2 showed that within a petrochemical company a whole department is responsible for managing the training and certification of all personnel. Thus, probably in hospitals too, dedicated teams, committees or departments are required to realize the execution of cycles for training and certification. But not only in hospitals should equipment have a more prominent role, also within skills courses it deserves more attention, as shown in chapter 4.2. Skills courses are meant to train the practice of (new) technologies and it is self-evident that the accompanying equipment is an essential part of these technologies and that these should be discussed sufficiently. Also within these courses there should be assessment and competence-based certification.

Another point of discussion is the target group for equipment training. Fundamentally, this is a matter of responsibility. According to national guidelines and the European law, whoever is required to use the equipment during his job, should be demonstrable competent in handling this equipment.¹¹⁻¹⁴ In practice, this concerns both the surgeon and the operating room (OR) nurse. Consequently, both should be trained in safe handling of equipment. Being an expert surgeon does not automatically imply that one is competent in this respect. In chapter 4.2 of this thesis a group of experts was asked to perform the same theoretical knowledge test on equipment as the course participants who were studied. This group of experts scored an average of 75%. It could be debated whether this score is acceptable for this group. Chapter 3.2 showed that regular training of fully specialized operators is standard practice in the petrochemical industry. Likewise, pilots have to repeat simulator trainings regularly to maintain their permission to fly an aircraft. Therefore, it seems plausible to suggest that anyone who uses equipment in the operating room, including a medical specialist, should train for this with regular repetition and especially demonstrate to be competent.

A drawback of training is that it takes time and effort. As described in the introduction of this thesis, the time of surgical residents currently spent in the OR is decreased compared to earlier times. Consequently, new training obligations are likely to reduce their OR time even further and thus training methods need to be efficient. Clinical lectures are time consuming and do not need to be attended live to reach an equal effect.¹⁵ Moreover, more active methods of training are more effective.¹⁶ Such other methods are at hand. In anesthesiology for instance, examples exist of scenario-based training with remote controlled equipment.¹⁷ The questionnaire used in chapter 4.1 also rendered some unique examples of hospitals where training of OR nurses takes place in small groups in which situations of equipment dysfunction are simulated. Such training takes more effort to organize but is likely to be more effective as it requires an active input of participants. On the other hand, a drawback of these methods is that it inflicts on clinical time. However, this can be obviated by methods for independent training. The web-based training modules described in chapters 4.3 and 4.4 are examples of such methods as they can be accessed from any computer and be performed in ones own time. The ultimate way to use such methods would be a training system with a database of all sorts of equipment from which the user can choose which device to train with. To achieve this, intensive cooperation is required between companies or institutions which develop such modules, medical associations and the industry which produces the devices. Currently, however, many device producing companies are focusing on development of their own training modalities for their own equipment and due to commercial interests these companies are not very open to cooperation. Consequently, different parties are individually going through the same process of development. Cooperation in, for instance, one united foundation would accelerate the development process and thereby save valuable time and money.

Future research and development

The modules described in chapters 4.3 and 4.4 are a start of a potentially successful way to train users of operating equipment. Issues that are not clarified yet are the exact target groups of the modules and the optimal training exposure. Moreover, the capacity of the assessments within the modules to discriminate in different levels of performance has not been studied yet. Future research should focus on determination of cut-off values for scores for 'pass' or 'fail' and their sensitivity and specificity. Another interesting aspect of these modules, which is currently being tested, would be to integrate them within VR simulators, so that surgical skills and equipment handling can be trained at the same time. As with real operations, such a combination can be used to train multitasking scenarios. Perhaps, with some kind of 'multiplayer' modality, this could even lead to team training for surgeons and OR nurses.

Ultimately, the clinical relevance of the training modules should be studied in practice. This could be facilitated by so-called intelligent data monitoring systems which can track and record all data concerning equipment, such as maintenance intervals, check, use, errors etc. Such a system is currently being developed and could not only help in the execution of scientific research, but also with the performance of preoperative equipment checks, risk analysis and planning. Therefore, such systems may also be a useful addition to safety management systems, which have to be implemented in all Dutch hospitals by now.

REFERENCES

- 1. Inspectie voor de Gezondheidszorg. Risico's minimaal invasieve chirurgie onderschat: kwaliteitssysteem voor laparoscopische operaties ontbreekt. Den Haag. 2007.
- 2. Marriott J, Purdie H, Crossley J, Beard JD. Evaluation of procedure-based assessment for assessing trainees' skills in the operating theatre. Br J Surg. 2011;98(450-7.
- Werkgroep endoscopische chirurgie (WEC) van de Nederlandse Vereniging voor Gastrointestinale Chirurgie namens de Nederlandse Vereniging voor Heelkunde. Minimaal invasieve chirurgie. Plan van aanpak en beleid: Nederlandse Vereniging voor Endoscopische Chirurgie (NVEC). 2009.
- Scott DJ, Rege RV, Bergen PC, Guo WA, Laycock R, Tesfay ST, et al. Measuring operative performance after laparoscopic skills training: edited videotape versus direct observation. J Laparoendosc Adv Surg Tech A. 2000;10(4):183-90.
- 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;257(3):476-82.
- 6. Orde van Medisch Specialisten. Verantwoordelijkheid medisch specialist bij onderhoud en beheer van medische apparatuur. Leidraad. Utrecht. 2008.
- Nederlands Technische Afspraak (NTA) 8009:2007. Safety Management System for hospitals and institutions which deliver hospital care. Delft, The Netherlands: Nederlands Normalisatie Instituut (NEN); 2007.
- 8. Projectgroep MICADO (Minimaal Invasieve Chirurgie Adequaat Door Ondersteuning). Kwaliteitsborging van instrumenten en apparatuur gebruikt bij minimaal invasieve chirurgie. 2008. p. 1-27 www.nvec.nl.
- 9. Verdaasdonk EG, Stassen LP, van der Elst M, Karsten TM, Dankelman J. Problems with technical equipment during laparoscopic surgery. An observational study. Surg Endosc. 2007;21(2):275-9.
- 10. Inspectie voor de Gezondheidszorg (IGZ). Risico's van medische technologie onderschat. Den Haag: Inspectie voor de Gezondheidszorg (IGZ). 2008. p. 1-81.
- 11. Nederlandse Vereniging van Ziekenhuizen (NVZ), Nederlandse Federatie van Universitair Medische Centra (NFU), Revalidatie Nederland. Convenant veilige toepassing van medische technologie in het ziekenhuis. Utrecht, The Netherlands. 2011.
- 12. Nederlandse Vereniging voor Anesthesiologie (NVA), Nederlandse Vereniging voor Heelkunde (NVvH). Richtlijn Het Peroperatieve Traject. Utrecht. 2011. p. 1-37.
- 13. European Parliament and Council. Directive 2000/34/EC of 22 June 2000. 2000.p.41-5
- 14. European Council. Council recommendation on patient safety, including the prevention and control of healthcare associated infections. 2009
- Rosser JC, Herman B, Risucci DA, Murayama M, Rosser LE, Merrell RC. Effectiveness of a CD-ROM multimedia tutorial in transferring cognitive knowledge essential for laparoscopic skill training. Am J Surg. 2000;179(4):320-4.
- 16. NTL Institute for Applied Behavioral Sciences. [22-02-2013]; Available from: http://www.ntl.org.
- 17. Berge JA, Gramstad L, Jensen O. A training simulator for detecting equipment failure in the anaesthetic machine. Eur J Anaesthesiol. 1993;10(1):19-24.

ENGLISH SUMMARY

Towards safer laparoscopic surgery: training and assessment of surgical skills and equipment handling

ENGLISH SUMMARY

Publication of the report "*To Err is Human*" has drawn attention to medical errors (**Chap-ter 1**). Worldwide studies have estimated that every 1 out of 10 patients suffers from an adverse event and that more than 40% of these events are preventable. An equal part of all adverse events takes place in or around the operating room and therefore preventive measures within the operative process are thought to have a great impact on patient safety.

Laparoscopic surgery has received special attention with regard to errors, as unusual complications were reported. It requires fundamentally different skills compared to conventional open surgery and it is heavily dependent on technology. Numerous reports have been published on laparoscopic surgery, which have led to several demands by authorities. It appeared that training for this operative technique was unstructured and not uniform, that there was no system for assurance of competence and there was a lack of training with the required technical equipment and no standardization of the whole operative process. In a reaction to this the Dutch Healthcare Inspectorate (DHI) has demanded structured training and implementation thereof into surgical training programs. Likewise, a system for intercollegiate objective assessment of surgical skills, preferably by video recordings, and use of means for standardization, like checklists, were recommended.

Presently, different methods for laparoscopic skills training are available, among which virtual reality simulators. These are most often used for basic skills training, however the exact borderline with other training modalities remains vague and so it is not clear what skills are to be trained with which method. For the objective assessment of surgical skills different methods exist as well, however it is not yet clear if one of these methods is suitable to be used for examination and certification, as is asked for by the DHI. Equipment-related safety seems to remain unaddressed, although incidents with equipment are not infrequent. Although means such as specific checklists for equipment have been shown to decrease the number of incidents, it is unknown whether these are actually being used in hospitals. Moreover, training with equipment seems to be taken less seriously than surgical skills training and not much is known about its effectiveness.

The objective of this thesis is to improve laparoscopic surgical skills training, with an emphasis on equipment-related safety and competence.

The first part of this thesis focuses on training and assessment of laparoscopic surgical skills. **Chapter 2.1** gives an extensive overview of all available methods for objective assessment of technical surgical skills and their evidence. This review of literature focuses on validity and reliability of methods used in general surgery and gynaecology. In total 104 studies were included, addressing 8 different categories of methods for assessment.

Of all these studies only 20 consisted of a high level of evidence. One of the most studied methods is the Objective Structured Assessment of Technical Skills (OSATS). This method was originally designed for use in a laboratory setting, but in several countries is now used as the official method for assessment of surgical residents in the operating room. However, only a minority of the validating studies addressed the use of OSATS in the operating room and none of these consisted of a high level of evidence. This chapter concludes that currently there is no method for objective assessment of technical surgical skills which is suitable to be used for summative assessment in the operating room, i.e. certification.

Chapter 2.2 focuses on virtual reality simulation and its position in training and assessment of more advanced laparoscopic skills. The study describes the face and construct validation of a curriculum for intermediate laparoscopic skills. This curriculum consists of 5 exercises which are all aimed at coordination between both hands and precision. All 5 exercises showed discriminative capacity, more or less. Thereby this chapter shows that not only basic laparoscopic skills, but also more advanced skills are demonstrable in a virtual reality environment. This suggests that training these skills with VR simulation might be feasible as well.

In the second part of this thesis healthcare is compared with another high-risk industry, the petrochemical industry, in order to filter out important issues for the improvement of safety. **Chapter 3.1** focuses on 'safety management systems' and the essence of a safety culture. A safety management system consists of systems for risk inventory and incident reporting to get an overview of the hazards that are threatening an organisation. For the optimal functioning of a safety management system, a high level safety culture is essential as risk inventory and incident reporting are fed by the organisation's employees. Currently, both safety management systems and safety culture are still in an immature stage in healthcare when compared to the petrochemical industry.

Chapter 3.2 focuses on training and assessment of equipment-related competence of personnel, referring to the increasing influence of technology in healthcare and especially surgery. A comparison is described of the management of equipment-related competence within a large petrochemical company and a large teaching hospital. In the petrochemical company competence management is arranged very strictly with defined competencies for every function and accompanying mandatory training cycles including assessment, certification and registration. In the studied hospital however, this is much more open-ended with no mandatory training, no assessment and no certification of personnel. Laparoscopic equipment is an exceptional category, for which in the studied hospital a yearly training cycle exists for OR nurses, since this was demanded by the healthcare inspectorate. In addition to the second part of this thesis, the third part emphasizes training with regard to technical laparoscopic equipment. First, an overview is given of the use of preoperative checklists for technical laparoscopic equipment and training of personnel with this equipment in Dutch hospitals. **Chapter 4.1** presents the result of a telephone-based questionnaire. The results of this chapter show that in all hospitals preoperative checklist. In the other hospitals a checking procedure is completed, but this is not standardized by the use of a checklist. The degree of detail among the checklists differs, illustrated by 9 checklists on which devices were not mentioned separately and almost half of the checklists on which connections was not a checkpoint. Finally, training with the equipment by OR nurses takes place in all hospitals, but the frequency is highly variable. Moreover, in one third of all hospitals training is not mandatory and only in 7 hospitals training includes an assessment. In almost all hospitals training with equipment is also deliberately aimed at residents and medical specialists.

To get an impression of equipment-related training for residents, in **chapter 4.2** the effect of 3 basic laparoscopic skills courses on knowledge of equipment is studied. These courses are aimed at young surgical residents at the beginning of their career, who are generally allowed to start practising laparoscopic surgery as an attending surgeon after completion of these courses. Therefore, it may be expected that part of what they learn during these courses addresses the equipment. However, the results show that the participants' knowledge of equipment only increased by 8% during these courses. Analysis of the course contents revealed that 2 of the 3 studied courses covered 50% or less of information which was considered essential. Moreover, 92.5% of the participants indicated that they would like to receive more training about equipment handling. Therefore, improvement of the contents of laparoscopic courses regarding equipment is advised.

In **chapters 4.3 and 4.4** the development and evaluation of 2 interactive web-based modules for surgical equipment is described. The module in **chapter 4.3** covers installation of a pneumoperitoneum and use of a laparoscopic insufflator and the module in **chapter 4.4** covers electrosurgery and the use of an electrosurgical device. Both modules consist of 3 parts: a theoretical part about the underlying physics, a tutorial in which is explained how exactly to use the device, and finally an assessment which tests both theoretical and practical competence. Both modules make use of a simulation of the devices in which connections can be established, buttons can be pushed and alarms can go off. These simulations are used for practice and assessment of practical competence. For both modules the training effect was studied for theoretical and practical competence. Chapters 4.3 and 4.4 show that training with both modules improves theoretical knowledge and practical skills significantly. In a test with the real devices

participants who trained with the modules made more than 50% less errors in operating the devices than participants who did not train with the modules. Moreover, participants who trained with the modules rated it with fairly good marks for user-friendliness, level of difficulty, training capacity and especially level of realism. All participants indicated the modules to be a useful addition to surgical training programs. These modules can be promising means to train healthcare personnel in the future. Main advantages are that they can be used by employees in their own time and without the need of an instructor.

In **chapter 5** general conclusions are drawn from this thesis, followed by a discussion and perspectives for future research.

NEDERLANDSE SAMENVATTING

Naar veiligere laparoscopische chirurgie: training en beoordeling van chirurgische vaardigheden en omgang met apparatuur

NEDERLANDSE SAMENVATTING

Met de publicatie van het rapport *"To Err is Human"* werd de aandacht gevestigd op medische fouten (**Hoofdstuk 1**). Wereldwijde studies schatten dat 1 op de 10 patiënten te maken krijgt met een medische fout en dat meer dan 40% van deze fouten te voorkomen zijn. Een even groot gedeelte van alle fouten vindt plaats rond of in de operatiekamer en daarom is de gedachte dat preventieve maatregelen binnen het operatieve proces een grote invloed zullen hebben op de patiëntveiligheid.

Met betrekking tot medische fouten is er bijzondere belangstelling voor laparoscopische chirurgie, omdat er bij deze vorm van chirurgie ongebruikelijke complicaties zijn gemeld. Deze vorm van minimaal invasieve chirurgie vereist andere vaardigheden dan conventionele open chirurgie en is veel meer afhankelijk van technologie. Er zijn meerdere rapporten over laparoscopische chirurgie verschenen welke hebben geleid tot verscheidene veiligheidseisen. Het bleek dat training voor deze operatietechniek niet gestructureerd was en niet uniform, dat een systeem voor het zekeren van bekwaamheid ontbrak, dat er een gebrek was aan training met de benodigde apparatuur en dat het gehele operatieve proces niet gestandaardiseerd was. Als reactie hierop eiste de Inspectie voor de Gezondheidszorg (IGZ) gestructureerde training en implementatie daarvan in de chirurgische opleidingen. Daarnaast werd er aangedrongen op een systeem voor intercollegiale beoordeling van chirurgische vaardigheden, liefst door middel van video opnames, en gebruik van middelen voor standaardisatie zoals checklijsten.

Momenteel bestaan er verschillende methoden voor het trainen van laparoscopische chirurgische vaardigheden, waaronder virtual reality simulators. Deze worden vaak gebruikt voor het trainen van basale laparoscopische vaardigheden. De grens tussen deze en andere trainingsmethoden blijft echter vaag en het is daarom niet duidelijk welke methode precies gebruikt kan worden voor het trainen van verschillende niveaus van vaardigheden. Ook voor objectieve beoordeling van chirurgische vaardigheden bestaan verschillende methoden. Het is echter nog niet duidelijk of een van deze methoden geschikt is als instrument voor examinering en/of certificering, zoals gewenst door de IGZ. Veiligheid gerelateerd aan operatieapparatuur blijft onderbelicht terwijl incidenten met deze apparatuur frequent voorkomen. Het is aangetoond dat middelen, zoals checklijsten speciaal gericht op apparatuur, het aantal incidenten met apparatuur kunnen verminderen. Het is echter niet bekend in hoeverre dit soort middelen ook daadwerkelijk worden gebruikt in ziekenhuizen. Bovendien lijkt het zo te zijn dat trainen met apparatuur minder serieus wordt genomen dan trainen op chirurgische vaardigheden. Er is ook niet veel bekend over de effectiviteit van trainingen met apparatuur.

Het doel van dit proefschrift is om training met betrekking tot laparoscopische vaardigheden te verbeteren, met een nadruk op de veiligheid en bekwaamheid gerelateerd aan de operatieapparatuur. Het eerste deel van dit proefschrift concentreert zich op trainen en beoordelen van laparoscopisch chirurgische vaardigheden. Hoofdstuk 2.1 geeft een uitgebreid overzicht van alle beschikbare methoden voor objectieve beoordeling van chirurgische vaardigheden. Dit overzicht van de literatuur concentreert zich op validiteit en betrouwbaarheid van methoden welke gebruikt worden in de chirurgie en gynaecologie. In totaal werden 104 studies geïncludeerd welke betrekking hadden op 8 verschillende categorieën van beoordelingsmethoden. Van alle studies hadden er slechts 20 een hoge bewijsgraad. Een van de meest bestudeerde methoden betreft de Objective Structured Assessment of Technical Skills (OSATS). Deze methode is ontworpen voor gebruik in een laboratoriumomgeving, maar wordt in verschillende landen nu gebruikt als officiële methode voor beoordeling van chirurgen in opleiding op de operatiekamer. Een beperkt aantal van de studies over deze methode heeft echter betrekking op zijn gebruik in de operatiekamer en geen van deze studies had een hoge bewijsgraad. Dit hoofdstuk concludeert dan ook dat er tot op heden geen methode voor objectieve beoordeling van chirurgische vaardigheden geschikt is voor examinering en/of certificering welke op de operatiekamer gebruikt kan worden.

Hoofdstuk 2.2 betreft de positie van virtual reality simulatie in de training en beoordeling van meer geavanceerde laparoscopische vaardigheden. Deze studie beschrijft de face- en construct validatie van een curriculum voor 'intermediate laparoscopic skills'. Dit curriculum bestaat uit 5 oefeningen welke allen gericht zijn op precisie en de coördinatie tussen 2 handen. Alle 5 oefeningen hadden in min of meerdere mate een onderscheidend vermogen in vaardigheid. Daarmee toont dit hoofdstuk aan dat niet alleen basale laparoscopische vaardigheden, maar ook meer geavanceerde laparoscopisch vaardigheden aantoonbaar zijn in een virtual reality omgeving. Dit suggereert dat het mogelijk is om ook dit soort vaardigheden te trainen met virtual reality simulatie.

In het tweede deel van dit proefschrift wordt de gezondheidszorg vergeleken met een andere 'hoog risico' sector, de petrochemische industrie, met als doel belangrijke onderwerpen voor het verbeteren van veiligheid te identificeren. **Hoofdstuk 3.1** beschrijft veiligheidsmanagement systemen en het belang van een veiligheidscultuur. Een veiligheidsmanagement systeem bestaat uit een systeem voor risico-inventarisatie en incidentmelding om zo een overzicht te krijgen van alle potentiële gevaren binnen een organisatie. Voor het optimaal functioneren van een veiligheidsmanagement systeem is een veiligheidscultuur van hoog niveau essentieel omdat de systemen voor risico-inventarisatie en incidentmelding afhankelijk zijn van de input van werknemers van een bedrijf. In de gezondheidszorg zijn veiligheidsmanagement systemen en een veiligheidscultuur nog in een onvolwassen stadium, vergeleken met de petrochemische industrie. **Hoofdstuk 3.2** beschrijft de vergelijking tussen een ziekenhuis en een petrochemisch bedrijf en beperkt zich daarbij tot de training en beoordeling van personeel met betrekking tot vaardigheid in apparatuurgebruik. In het petrochemisch bedrijf is dit zeer strikt georganiseerd met gedefinieerde vaardigheden voor elke functie en bijbehorende trainingcycli, beoordeling, certificering en registratie daarvan. In het bestudeerde ziekenhuis was dit veel minder strikt georganiseerd zonder verplichte trainingen, beoordelingen en certificatie van personeel. Laparoscopische operatieapparatuur bleek een uitzondering, want daarvoor bestond een jaarlijkse trainingscyclus voor operatieassistenten, omdat dit een eis was van de IGZ.

In aansluiting op het tweede deel van dit proefschrift, wordt in het derde deel de nadruk gelegd op training met laparoscopische operatieapparatuur. Allereerst wordt een overzicht gegeven van het gebruik van checklijsten voor laparoscopisch operatieapparatuur en de training van personeel met deze apparatuur in Nederlandse ziekenhuizen. In Hoofdstuk 4.1 worden de resultaten beschreven van een telefonische enquête. De resultaten van dit hoofdstuk laten zien dat in alle ondervraagde ziekenhuizen preoperatieve controle van laparoscopische operatieapparatuur plaatsvindt en dat dit in 56% van de ziekenhuizen gebeurt met behulp van een checklist. In de andere ziekenhuizen wordt wel een preoperatieve controle uitgevoerd, maar gebeurt niet gestandaardiseerd aan de hand van een checklist. De mate van detail van de checklijsten verschilt, hetgeen blijkt uit het feit dat op 9 checklijsten apparaten niet afzonderlijk werden benoemd en op bijna de helft van alle checklijsten de aansluitingen geen punt van controle was. Tot slot wordt in bijna alle ziekenhuizen getraind met operatieapparatuur door operatieassistenten, maar de frequentie daarvan is hoogst variabel. Bovendien is in een derde van de ondervraagde ziekenhuizen trainen niet verplicht en zijn er slechts 7 ziekenhuizen waar vaardigheden ook beoordeeld worden. In vrijwel alle ziekenhuizen was training met apparatuur gericht op de operatieassistenten. Slechts in 4 ziekenhuizen was de training ook specifiek gericht op arts-assistenten en medisch specialisten.

Om een indruk te krijgen van apparatuur gerelateerde training voor arts-assistenten, wordt in **hoofdstuk 4.2** het effect bestudeerd van 3 cursussen voor basale laparoscopische vaardigheden. Deze cursussen zijn gericht op jonge arts-assistenten die aan het begin van hun carrière staan. Over het algemeen mogen zij na deze cursus als eerste operateur fungeren bij laparoscopische operaties. Het mag derhalve verwacht worden dat apparatuur een van de onderwerpen is tijdens deze cursussen. De resultaten van deze studie laten echter zien dat deze cursussen het kennisniveau van de kandidaten slechts met 8% laten stijgen. Analyse van de inhoud van de cursussen toonde dat in 2 van de 3 cursussen minder dan 50% van de essentieel geachte informatie over apparatuur aan bod kwam. Bovendien gaf 92,5% van de deelnemers van deze cursussen aan

meer training te willen in bediening van de apparatuur. Aanpassing van de inhoud van basis laparoscopische cursussen lijkt dus geïndiceerd.

In hoofdstuk 4.3 en 4.4 wordt de ontwikkeling en evaluatie van 2 interactieve webbased trainingsmodules voor laparoscopische operatieapparatuur beschreven. De module beschreven in **hoofdstuk 4.3** betreft de installatie van een pneumoperitoneum en het gebruik van een insufflator. De module in **hoofdstuk 4.4** betreft elektrochirurgie en het gebruik van een elektrochirurgisch apparaat. Beide modules bestaan uit 3 delen: een theoretisch gedeelte over de onderliggende natuurkundige principes, een tutorial waarin precies wordt uitgelegd hoe het apparaat bediend moet worden en tot slot een assessment waarin zowel theoretische als praktische vaardigheid getest worden. In beide modules wordt gebruik gemaakt van simulatie van het daadwerkelijke apparaat waarbij aansluitingen kunnen worden gemaakt, knoppen kunnen worden bediend en alarmen af gaan. Deze simulaties worden gebruikt voor oefening en voor beoordeling van praktische vaardigheid. Hoofdstuk 4.3 en 4.4 laten zien dat trainen met beide modules de theoretische en praktische vaardigheden significant verbetert. In een test met de daadwerkelijke apparaten maakten deelnemers die met de modules getraind hadden ruim 50% minder fouten in de bediening van de apparaten dan deelnemers die niet hadden getraind. Bovendien werden de modules door de deelnemers goed beoordeeld met betrekking tot gebruiksvriendelijkheid, moeilijkheidsgraad, trainingscapaciteit en vooral het realisme. Alle deelnemers gaven aan dat de modules een nuttige toevoeging zouden zijn aan chirurgische opleidingen. Deze modules zijn een veelbelovende trainingsmethode voor personeel in de gezondheidszorg. De voornaamste voordelen zijn dat ze zelfstandig gebruikt kunnen worden in de eigen tijd en zonder de benodigdheid van begeleiding.

In **hoofdstuk 5** worden algemene conclusies getrokken, gevolgd door een discussie en toekomstperspectieven voor onderzoek.

LIST OF PUBLICATIONS

- van Hove PD, Tuijthof GJ, Verdaasdonk EG, Stassen LP, Dankelman J.
 Objective assessment of technical surgical skills.
 Br J Surg. 2010 Jul;97(7):972-87. doi: 10.1002/bjs.7115. Review.
- van Hove PD, Verdaasdonk EG, Dankelman J, Stassen LP. Hoe SCHERP is de OSATS? Nederlands Tijdschrift voor Heelkunde. 2010 Dec;19(9): 343-347.
- Schreuder HW, van Hove PD, Janse JA, Verheijen RR, Stassen LP, Dankelman J. An "intermediate curriculum" for advanced laparoscopic skills training. *J Minim Invasive Gynecol.* 2011 sep-oct;18(5):597-606. doi: 10.1016/j.jmig.2011.05.017
- **van Hove PD**, Verdaasdonk EG, Van der Harst E, Jansen FW, Dankelman J, Stassen LP. Effect of basic laparoscopic skills courses on essential knowledge of equipment. *Surg Innov.* 2012 Dec;19(4):460-6. doi: 10.1077/1553350611432723.
- van Hove PD, Verdaasdonk EG, Dankelman J, Stassen LP.
 Gebruik van checklists voor laparoscopische apparatuur in Nederlandse ziekenhuizen.
 Ned Tijdschr Geneeskd. 2012;156(33):A4505. Dutch.
- van Hove PD, van Wijk R, Verdaasdonk EG, Stassen LP, Dankelman J. Training and assessment of equipment-related competence. Comparison of a petrochemical company and a hospital. *Health Technol.* 2013 Feb. Epub ahead of print.

LIST OF CONGRESS PRESENTATIONS

- van Hove PD, Tuijthof GJM, Verdaasdonk EGG, Stassen LPS, Dankelman J. Methods for Objective Assessment of Technical Surgical Skills.
 21st Conference of the Society for Medical Innovation and Technology (SMIT). 2009, Sinaia, Romania.
 Oral presentation.
- van Hove PD, Tuijthof GJM, Verdaasdonk EGG, Dankelman J, Stassen LPS. Hoe SCHERP is de OSATS? Najaarsvergadering 2009, Nederlandse Vereniging voor Heelkunde (NVVH). Within: Symposium "Het nieuwe opleiden". 2009, Ede, the Netherlands. Oral presentation.
- van Hove PD, Schreuder HWR, Janse JA, Dankelman J, Stassen LPS.
 Virtual Reality Training voor de Gevorderde Laparoscopist..
 8th congress of the Nederlandse Vereniging voor Endoscopische Chirurgie (NVEC). 2010, Amersfoort, the Netherlands.
 Poster presentation.
- van Hove PD, Schreuder HWR, Janse JA, Dankelman J, Stassen LPS.
 Virtual Reality Training voor de Gevorderde Laparoscopist..
 Chirurgendagen 2010, Nederlandse Vereniging voor Heelkunde (NVVH). 2010, Veldhoven, the Netherlands.
 Poster presentation.
- van Hove PD, Schreuder HWR, Janse JA, Stassen LPS, Dankelman J.
 Face and construct validity of a virtual reality curriculum for intermediate laparoscopic surgeons.
 18th conference of the European Association of Endoscopic Surgery (E.A.E.S.). 2010, Geneva, Switzerland.
 Oral presentation.
- van Hove PD, Verdaasdonk EGG, van der Harst E, Jansen FW, Dankelman J, Stassen LPS.

Training for equipment handling. For all OR staff?

3rd congress of the Dutch Society for Simulation in Healthcare (DSSH). 2011, Utrecht, the Netherlands.

Oral presentation.

• van Hove PD

Simulatietraining met medical devices. 4th congress of the Dutch Society for Simulation in Healthcare (DSSH). 2012, Utrecht, the Netherlands. Workshop.

- van Hove PD, Verdaasdonk EGG, Dankelman J, Stassen LPS. Training laparoscopic equipment handling with an interactive simulation module. 20th conference of the European Association of Endoscopic Surgery (E.A.E.S.). 2012, Brussels, Belgium. Oral presentation.
- van Hove PD, Verdaasdonk EGG, Dankelman J, Stassen LPS.
 Web-based training of equipment use.
 21st conference of the European Association of Endoscopic Surgery (E.A.E.S.), 2013,
 Vienna, Austria.
 Oral presentation.

CURRICULUM VITAE

Diederick van Hove was born on the 31st of May 1983 in The Hague, the Netherlands. After the "Montessori school Waalsdorp" he attended the "Maerlant Lyceum" in The Hague and graduated in 2001. In the same year he started his medical study at the University of Leiden in Leiden, the Netherlands. During his study he worked as an administrative employee at the first aid department of the Leiden University Medical Center. For a scientific internship Diederick performed a retrospective study on calcaneal fractures together with trauma surgeon K.A. Bartlema at the Leiden University Medical Center. At the end of his clinical internships he attended his final internship at the department of surgery in the Haaglanden Medical Center (MC Haaglanden) at the location 'Westeinde' in The Hague, the Netherlands (dr. J.C.A. de Mol van Otterloo).

Diederick obtained his medical degree on the 29th of May 2008. After this he started his medical career as a surgical resident (not in specialist training) at the department of surgery of the Reinier de Graaf Gasthuis in Delft, the Netherlands (dr. L.P.S. Stassen). In January 2009 he started his PhD research at the department of Biomechanical Engineering of Delft University of Technology in Delft, the Netherlands, under the supervision of prof. J. Dankelman and prof. L.P.S. Stassen.

From 2009 to present Diederick has organized simulator training with regard to the Rotterdam basic laparoscopic skills course, twice every year. In 2010 he was a member of the organizing committee of the "Symposium Experimenteel Onderzoek Heelkundige Specialismen" (SEOHS) in Rotterdam.

In January 2012 Diederick resumed his work as a surgical resident (not in specialist training) at the department of surgery of the Maasstad Hospital in Rotterdam, the Netherlands (dr. E. van der Harst). He currently holds this position.

DANKWOORD

Onderzoek doen en een proefschrift schrijven doe je niet alleen. Er is een aantal mensen zonder wie dit proefschrift zonder meer niet tot stand was gekomen, of die op een andere manier zeer van belang zijn geweest tijdens mijn onderzoeksperiode. Deze mensen verdienen het om op deze plaats genoemd te worden.

Prof. Dr. J. Dankelman, beste Jenny,

Mijn ervaringen als onderzoeker aan de TU zullen voor de rest van mijn werkzame leven als arts van waarde zijn. Van jou heb ik geleerd om op een andere manier naar processen en problemen in de zorg te kijken; iets wat mijns inziens voor elke arts van waarde zou zijn. De samenwerking met jou was altijd gemakkelijk en laagdrempelig, want jouw deur stond letterlijk altijd open. Sinds jouw voorzitterschap van de afdeling BME, had je duidelijk minder tijd over, maar als het nodig was dan maakte jij altijd tijd. Ook heb jij altijd rekening gehouden en meegedacht met mijn verdere carrière. Dat heb ik enorm gewaardeerd. Bedankt voor alles!

Prof. Dr. L.P.S. Stassen, beste Laurents,

Jou ben ik misschien wel de meeste dank verschuldigd van iedereen. Jij zag in mij de geschikte kandidaat voor voortzetting van het onderzoek aan de TU en jij bood mij (samen met Jenny) deze kans. Ik denk dat weinig onderzoekers een (co-)promotor hebben die zo goed, tot in elk detail, op de hoogte is van hun werk en zich daar zo grenzeloos voor inzet. Ook als dingen wat minder gingen had jij daar alle begrip voor. Jouw verhuizing naar Maastricht heeft mij een kort moment doen twijfelen over het voortbestaan van onze samenwerking, maar het is bewezen dat dit daar allerminst aan heeft afgedaan. Vaak genoeg ben jij afgereisd naar Delft om daar met mij de lopende zaken face-to-face te bespreken. Andersom ben ik ook veel in Maastricht geweest en ik heb me daar altijd zeer welkom gevoeld danzij jouw gastvrijheid. Enorm veel dank voor deze kans, het vertrouwen, en de sublieme samenwerking! Het resultaat ligt voor je.

Dr. E.G.G. Verdaasdonk, beste Emiel,

Tijdens een gezamenlijke dienst op de SEH in het Reinier de Graaf vertelde jij me over jouw promotieonderzoek aan de TU Delft en een nieuwe vacature aldaar. Ik ben je daar nog steeds dankbaar voor. Jij hebt altijd een onmiskenbare rol gespeeld in mijn onderzoek. Ik kan een koppig persoon zijn en onze meningen lagen soms ver uit elkaar, maar jouw commentaar heeft mijn werk altijd positief beïnvloed. Je hebt daar zeer veel tijd in gestoken en ik ben je dan ook zeer dankbaar voor al die vrije uren die jij hebt besteed aan grondige revisies van mijn stukken. Ze zijn er allen beter door geworden. Dank, dank en nog eens dank! De overige leden van mijn promotiecommissie: Prof. Dr. M. Mulder, Prof. Dr. F. Scheele en Prof. Dr. J.F. Lange,

Ik dank u allen hartelijk dat de tijd heeft genomen om mijn manuscript te beoordelen en dat u zitting hebt willen nemen in mijn promotiecommissie om mij het vuur aan de schenen te leggen tijdens de verdediging.

Prof. T.A Nagelhus Hernes,

Thank you very much for the willingness to review my manuscript and to come to Delft to take place in the opposing committee at the defence of my thesis.

Alle co-auteurs van de artikelen: Dr. Gabrielle Tuijthof, Dr. Henk Schreuder, Drs. Julienne Janse, Prof. Frank Willem Jansen, Dr. Erwin van der Harst en Dr. Ron van Wijk,

Niet alleen hebben jullie meegeschreven aan de artikelen, maar jullie hebben ook allemaal een onmisbare bijdrage geleverd aan de onderzoeken die daaraan vooraf gingen. Gabrielle, samen hebben we meer dan 100 abstracts doorgespit. Dat was veel werk, maar we hebben er een goed review over geschreven. Henk en Julienne, door de perfecte samenwerking hebben we een groot aantal proefpersonen kunnen testen op de Simendo. Prof. Jansen en Dr. van der Harst, dank voor de mogelijkheid om een deel van mijn onderzoek uit te voeren tijdens jullie laparoscopische cursussen. Ron van Wijk, dank voor de toegang en de contacten bij Shell. Ik heb daar erg veel van geleerd.

De mannen van Simendo: Joeri, Mark, Erwin en Michael. Zonder jullie medewerking was dit een heel ander proefschrift geworden. Een idee is namelijk leuk, maar er moet wel iemand zijn die het kan uitvoeren. Jullie zagen daar wel wat in.

Joeri, ik dank je voor het vertrouwen dat jij in mij gesteld hebt door je met Simendo (geheel kosteloos) voor 100% in te zetten voor het ontwikkelen van de modules voor apparatuur. Dit is tekenend voor jullie mooie bedrijf.

Dr. J.J. van den Dobbelsteen, beste John,

Ook al hebben wij nooit echt samen gewerkt, als ik hulp nodig had dan kon ik altijd bij jou aankloppen voor een pragmatisch advies. Als psycholoog tussen de ingenieurs begreep jij dat dat voor mij, als arts, ook even wennen was. Bedankt!

Alle (oud-)collega's van de TU: Arjo, Jasper, Winfred, Mehdi, Helene, Gerwin, Sander, Tim, Dennis, Kirsten, Anne, Paul, Bart, en natuurlijk Anouk, Diones en Dineke (en iedereen die ik nog vergeet te noemen).

Bedankt dat jullie mij als 'niet-ingenieur' al die jaren hebben gedoogd. Van en met jullie heb ik de kneepjes van het onderzoeken geleerd. Ik wens jullie allen veel succes bij het afronden van jullie promoties (voor wie dat nog moet doen) en bij jullie verdere loopbanen.

Alle stafleden en arts-assistenten van de afdeling Chirurgie van het Reinier de Graaf Gasthuis in Delft waar ik mee heb samen gewerkt,

Bedankt voor de leuke tijd die ik bij jullie gehad heb en voor de interesse voor en medewerking aan mijn onderzoek. Ik heb mij vanuit Delft altijd gesteund gevoeld!

Alle stafleden en arts-assistenten van de afdeling Chirurgie van het Maasstadziekenhuis in Rotterdam,

Ik heb in de afgelopen anderhalf jaar met veel plezier enorm veel van jullie mogen leren. Bedankt daarvoor, maar ook voor de interesse in mijn onderzoek en de flexibiliteit die geboden is bij de afronding ervan. Vooral ook bedankt aan diegenen die, vaak na werktijd, als proefpersoon hebben willen fungeren!

De mannen van GeVaT, mijn cordial: Bob, Mark, Onno, Luuc, Exie pexie, Muis, Eljee, Uljee, Mud, GP, Jay, Spoel, van Melle en IJs,

Gasten, dank voor alle mooie, gezellige en vaak ook beschonken avonden en feestjes waarvan er vast nog velen zullen volgen. Ze zijn een goede afleiding geweest. Ik ben er trots op dat wij nu al 11 jaar zo'n bijzonder hecht cordial zijn en elkaar nog zo regelmatig zien. Dalen, maar vooral ook toppen hebben we met zijn allen meegemaakt. Ik schaar dit proefschrift onder de toppen. Dank mannen!

Bob en Mark, mijn paranimfen, cordialgenoten, ex-huisgenoten, collega-doctors, maar vooral mijn zeer goede vrienden,

Dank dat jullie aan mijn zijde staan, niet alleen bij de verdediging van mijn proefschrift, maar ook de afgelopen 11 jaren dat wij elkaar al kennen. Bobbelaer, vaak genoeg heb ik tijdens mijn onderzoek dankbaar gebruik mogen maken van je statistische kennis en kunde als epidemioloog en 'wandelende SPSS-handleiding'. Met jullie beide heb ik regelmatig ervaringen als promovendus kunnen uitwisselen en dat was vaak leerzaam. Bedankt vrindjes!

Lieve Robert, Irene, Anne, Wouter, Willemijn, Dirk en Renske,

Ik overdrijf niet als ik zeg dat ik me geen betere schoonfamilie had kunnen wensen. Jullie vormen een bijzonder hecht en warm gezin en ik ben blij dat ik daar deel van mag uitmaken. De vele gezellige (en altijd goed gecaterde) avondjes zijn dan ook nooit een straf. Dank voor al jullie oprechte interesse en steun! Lieve pap en mam,

Bedankt voor... alles! Jullie hebben mij altijd mijn eigen weg laten gaan en daarin was niets onmogelijk. Dankzij jullie ben ik geworden tot wie ik nu ben en ik ben er van overtuigd dat dat er toe heeft bijgedragen dat dit proefschrift hier nu voor jullie ligt. Ook al was dat de afgelopen jaren niet makkelijk, allebei hebben jullie mij op je eigen manier gesteund. Ik hou van jullie.

Lieve broertjes en zusjes: Wouter, Jinne, Lilian en Gep,

Bedankt voor jullie interesse in mijn werk de afgelopen jaren, maar vooral bedankt voor alle relativerende lol en gekkigheid. Never a dull moment!

En dan *last*, maar zeker niet *least*:

Lieve lieve Femke,

Wat ben ik blij dat onze wegen elkaar ooit gekruist hebben. Dank voor al je oprechte interesse, steun en liefde de afgelopen jaren. Je zat me soms meer achter m'n broek dan mijn promotores ;-) Vooral de laatste anderhalf jaar, waarin mijn werk in het ziekenhuis gecombineerd moest worden met het afronden van het laatste onderzoek en het proefschrift, heb ik jou behoorlijk tekort moeten doen. Jij hebt daar nooit moeilijk over gedaan en het geaccepteerd zoals het kwam, maar dat is niet vanzelfsprekend. Binnenkort kunnen de weekenden en vakanties weer gebruikt worden waarvoor ze bedoeld zijn: samen genieten! Love you!

