General introduction
Surgical techniques and surgical education: from past to present

Over the past decades there has been a great leap forward in the evolution of surgical techniques and surgical education. Since minimal invasive surgical techniques (laparoscopy and endoscopy) were introduced around 1980, major benefits have been shown for the patient, such as reduced postoperative pain, shorter hospital stay, and cosmetic benefits. For the trainee however, the time to master these minimal invasive techniques is longer compared to open surgery. This is due to e.g. decreased haptic feedback, less range of motion of the instruments, and disparity between visual and proprioceptive feedback. While the time that is necessary to reach a predefined level of competency is longer, the residents’ clinical training hours have been reduced to an average of 48 hours per week due to the introduction of the European Working Time Directive in 2009. This has raised constraints in attaining appropriate levels of competency during residency training. Furthermore, it is becoming ethically unacceptable to perform a first procedure directly on the patient. In today’s legal environment more importance is placed on patient safety and the standards for clinicians’ competency are higher. All these developments have challenged the traditional master-apprentice type of training of ‘see one, do one, teach one’, and have led to several major changes within surgical education.

The first major change within surgical education is the onset of simulation-based training (SBT). The key advantage of SBT is that the initial phase of the residents’ surgical learning curve (in which the greatest progress is made) is moved to a risk-free and time-independent environment. Since the onset of SBT, a great number of simulation models have been designed, and several studies have shown that surgical skills can be improved by SBT. Moreover, SBT can improve patient comfort and reduce complication rates, thereby enhancing patient safety. Nowadays, SBT is becoming increasingly accepted as a method to complement training in clinical practice. In the United States, this has recently lead to a statement of the Residency Review Committee for Surgery mandating that ‘simulation and skills facilities must be available for all residents.’ In The Netherlands, a similar, though indirect, statement has been included in the framework decision (kaderbesluit) of the Central College of Medical Specialties, stating that: ‘enough instruments, spaces, and other facilities must be available to ensure a high quality education for the specialty in question’. The second major change within surgical education is the gradual shift from time-based training, which ends after a predefined time span, towards competency-based training, in which competency levels are defined, that have to be met before a resident is allowed to perform a certain procedure independently.
not threatening patient safety. In this type of training, objective assessment of skills is a prerequisite for final competency levels to be guaranteed.

With the gradual shift towards competency-based training including the rise of SBT, surgical education has entered a new phase. The current challenge for surgical educators is to integrate validated SBT and objective assessment of skills into the existing surgical curricula.

Simulation-based training

**Development, validation and implementation of SBT**

In the development of SBT, it is important to start with defining the intended and desired outcomes of the training. A *training needs analysis* (TNA) can be useful to identify training objectives, including procedural steps and pitfalls. The performance of a TNA consists of three major steps: (1) a description of general steps of the procedure based on expert opinion and international guidelines, (2) development of a general inventory of the most common (type and complexity) pitfalls of the procedure, and (3) a detailed analysis of pitfalls in real-time procedures on patients. After training needs and desired outcomes have been defined, a suitable simulator can be selected. There are numerous simulators available for the training of urological basic and procedural skills, ranging from low fidelity to high fidelity simulators. The *fidelity* of a simulator relates to its degree of realism. The choice for a low or high fidelity simulator depends on the target group, the aim of training, and the available budget.

Before a simulator is incorporated into urological residency training it needs to be validated, to ensure that residents acquire the required operative skills by training on the simulator. The *validation process* of a simulator entails a subjective and an objective approach. The subjective approach comprises face and content validation, defined as “the judgement of novices regarding realism and usefulness of the simulator” and “the judgement of experts regarding realism and usefulness of the simulator” respectively. The objective approach consists of the evaluation of construct and criterion validity. It is generally considered that the minimum evidence for widespread simulator use should include face, content, and construct validity. Construct validity relates to the degree a simulator can train or measure what it intends to train or measure. For example: the ability of a simulator to differentiate between different levels of experience. Criterion validity consists of concurrent validity and predictive validity. For the establishment of concurrent validity, performance on a simulator is correlated with a previously validated technique/gold standard. Predictive validity relates to the effect of simulation training on the actual performance in the operating room. Predictive validity cannot be confirmed without a valid and reliable assessment tool that enables the measurement of operative performance.
Although the benefits of SBT are apparent, training programs struggle with the integration of SBT into their curricula. This is due to varying opinions on the content of SBT, the validation process and the method of integrating SBT in a training curriculum. Moreover, issues such as considerable cost, logistics and limited personnel play a significant role.

**Theoretical framework in the design of SBT curricula**

There are several evidence-based educational theories, principles, and concepts that should be taken into consideration in the design of a simulation-based training curriculum. One of the prevailing educational theories on motor skill acquisition is the **three-stage theory of Fitts and Posner**, in which three sequential phases are described. The first phase is the cognitive phase, which focuses on the understanding of a complex task including its procedural steps. The second phase is the association phase, in which the mental process is put into practice and the complex task is trained. The final phase is the autonomous phase, in which the skill performance becomes a non-conscious automatic process. Ideally, basic urological skills should be trained to automaticity in a non-patient related setting, so that residents can focus on the performance of the actual procedure in the clinical setting.

Training to automaticity relates to the concept of **deliberate practice** as described by Ericsson: ‘to achieve excellent performance, residents need to train their skills over and over again’. According to Ericsson there are five conditions required for deliberate practice: a well-defined task to practice, a variable level of task difficulty, provision of instant, constructive feedback, motivation to improve, and frequent opportunity to repetition. The availability of these five conditions should be kept in mind in the design of simulation-based training. When studying the optimal interval of skills training, literature suggests that it would be better to train frequently for shorter periods instead of less frequently for longer periods. This relates to the principle of **distributed practice**. Residents retain and transfer skills better if they are trained in a schedule of distributed practice instead of massed practice. Finally, based on the concept of **whole-task training**, training should aim at learning the procedure from beginning to end, because ‘in clinical practice, the whole is much greater than the sum of its parts’.

The three aspects that should receive attention in whole-task training include the acquisition of cognitive knowledge, the acquisition of technical - psychomotor - skills, and the acquisition of non-technical skills (e.g. communication, decision-making, teamwork, and situational awareness). Literature has shown that adverse events in healthcare mostly occur due to medical errors, which in part are preventable. In the prevalence of medical errors, individual technical ability, knowledge and non-technical (human factor) skills play an important role. In this, human factor errors lead to most errors and adverse patient outcomes. In the design of an SBT curriculum it could be valuable to analyze the errors that occur during certain procedures, including the
(potential) harm they cause to the patient and how this can be prevented by training. Insight into the causes of adverse events can be reached by the performance of a root cause analysis (RCA). RCA exists in several forms and the one used in this thesis was the PRISMA-Medical version. Classification of the root causes is done in a predefined order, aiming to first explore technical and organisational factors before assigning a ‘human’ contribution to the incident in question.

Competency-based residency training

From time-based training towards competency-based training

As mentioned previously, the second major development within residency training is the gradual shift from time-based training, which ends after a predefined time span, towards competency-based training, which ends after predefined competency levels have been reached. In the Netherlands, there are two driving forces behind this shift. First, the Dutch government has announced extensive economic savings on postgraduate training. The budget for the residency training programs will be reduced with 218 million euros to be reached in 2022. One of the measures necessary to realize these savings, is a reduction of total residency training time with an average of 6 months. As quality has to be preserved, the concept of ‘individualized residency training’ is introduced. This means that residents with specific work experience or residents who reach proficiency earlier than others, will obtain a reduction in their training time.

Second, in the current Dutch legal environment an intensified focus is placed on patient safety. To ensure the highest quality of care, all physicians have to be ‘demonstrably qualified and competent’ in handling high-risk medical technology. This implicates the development of assessment tools that enable the measurement of proficiency in handling high-risk medical technology. The Dutch Health Care Inspectorate has recently performed an official evaluation on the integration of this assessment in Dutch hospitals. They concluded that this integration is still insufficient and that ongoing efforts to accomplish this goal have to be made.

Due to the shift towards competency-based training, benchmark levels of competency have to be set and assessed. But what are the different stages of clinical performance and when is a resident judged to be competent?

Theory of competency-based training

There are several educational concepts and models that have focused on competency-based training. The Dreyfus and Dreyfus model describes a framework for skill acquisition that consists of five developmental stages: novice, advanced beginner, competent, proficient, and expert. This model was originally applied to skills as playing chess and driving a car, but is now also used by medical educators in describing and ultimately assessing clinical competence of physicians as they progress from
novice to expert. In this model a competent learner is defined as “a person that has a good working and background understanding, sees actions at least partly in context and is able to complete work independently to a standard that is acceptable though it may lack refinement”. A proficient learner is defined as “a person that has a deep understanding, sees actions holistically, can achieve a high standard routinely”. In literature, there is a lack of uniformity regarding the definition of competence and how competence is differentiated from proficiency. In addition, various types of competency (technical, operative, surgical) are described and defined. In this thesis we will use the term surgical competence, defined as: “a collection of skill, knowledge, and judgment required to complete new or familiar tasks incorporating both technical and non-technical (i.e. clinical problem solving) components.”

**Formative and summative assessment in competency-based training**

Up till now, the majority of assessment tools have been used for formative assessment. In formative assessment, the focus is on development and progression of residents throughout their traineeship by providing a structured evaluation of their performance and constructive feedback. In summative assessment, the goal is to assess the residents’ performance at the end of a course by comparing it against some standard or benchmark. Summative assessment tools are used for high-stakes assessment and certification, which are imperative in competency-based training. For an assessment tool to be used for high-stakes assessment, it has to comply with stringent requirements, as it has to be objective, feasible, valid, and reliable.

The task specific checklist is a form of assessment that is mainly used for formative purposes. This checklist contains yes/no questions relating to specific steps or tasks of a procedure. The use of a task specific checklist mandates performance of the task in a predefined order. Consequently, failure in performing the procedure in this predefined order can result in a low score. Although the use of a task specific checklist provides residents with specific feedback, its rigid nature leads to poor validity and reliability when applied to experienced surgeons.

Another example is the global rating scale (GRS). The GRS often uses a Likert scale and focuses on global parameters of operative performance. The GRS tends to be more reliable than task specific checklists, but does not provide specific feedback regarding the parts of a procedure that requires most attention.

Finally, perhaps the most well-known and widely used assessment tool that is used for formative and summative assessment is the Objective Structured Assessment of Technical Skills (OSATS), in which technical skills are assessed in a simulated or clinical setting. In OSATS, the use of a task specific checklist is combined with a global rating scale. This offers formative feedback and provides a general judgement of performance.

In the Netherlands, the first step towards summative assessment in Urology was the development and validation of the Program for Laparoscopic Urological Skills.
This program was designed by Tjiam and colleagues in response to calls from healthcare authorities for well-defined proficiency standards to guarantee the quality of care and optimize patient safety. The program assesses basic laparoscopic urological skills of junior residents and has been implemented on a national scale. Further training programs and assessment tools for endoscopic and robot-assisted urological procedures are still work in progress, including their validation, i.e. effect on patient outcomes and thereby patient safety.\(^6^8\)

**Methodology for development of high-stakes assessment**

A constructive methodology that can be used in the development of a summative assessment tool that measures surgical competency is **cognitive task analysis** (CTA).\(^{44}\) CTA aims to capture knowledge, procedural steps, thought processes and decision making during the performance of a complex task to identify conscious and unconscious skills.\(^{69-71}\) As experts have automated parts of their performance, they are no longer conscious of every step they take and have difficulties in identifying these points where certain decisions are made.\(^{69,72,73}\) CTA offers a unique educational method to deconstruct the automated skills of experts and to identify relevant steps and decision points.\(^74\)

**GENERAL PROBLEM DEFINITION**

In the first two parts of this thesis, we focus on the design, content, implementation, and evaluation of a urological SBT curriculum for residents. In the third part, we focus on the development and validation of a summative assessment tool that measures surgical competence in transurethral resection of a bladder tumour (TURBT) in a validated simulation setting.

SBT has been proven to be effective, but its successful integration within surgical curricula is still uncommon.\(^{34,36}\) In the development of SBT, it is paramount that training needs are identified before the content of a training module is determined. Furthermore, the question arises what design and implementation strategy should be used in a simulation-based skills training curriculum to achieve acceptability and optimize learning outcomes and patient safety.

Due to the gradual shift towards competency-based training, attention towards objective assessment and certification of skills is growing to ensure the highest quality of care and patient safety. But when is a resident or surgeon considered to be competent? And what approach should be used in the development and validation of a summative assessment tool that measures surgical competency? In this thesis, we address these questions in relation to the TURBT procedure as the first endoscopic urological procedure for which a high-stakes assessment tool is developed. For this procedure
was chosen as it is a vital and technically challenging urological procedure that demands high technical accuracy in order to obtain an optimal oncological outcome.\textsuperscript{75}

The two main research questions of this thesis are: “What is the optimal design, content and implementation strategy for a urological simulation-based training curriculum and what is its educational impact?” and “How can competency in technical and non-technical skills of TURBT be assessed in a validated, standardized simulation setting?” To address the general research questions we formulated the following six specific research questions:

**RESEARCH QUESTIONS**

1. How well do junior and senior residents perform on the patient regarding four basic urological procedures, considering completeness of procedural steps, level of independence, intervention time and the incidence of unintended events?

2. How do patients experience diagnostic urological procedures performed by urologists, junior and senior residents? What is the influence of procedure related factors on patient experiences?

3. How do residents currently and ideally learn their practical urology skills? Which design characteristics may increase the acceptability of urological practical skills training curricula such as the Dutch Urological Practical Skills curriculum?

4. What is the perceived educational impact of the national implementation of the Dutch Urological Practical Skills curriculum? What are focus points for improvement of the Dutch Urological Practical Skills curriculum according to the participants?

5. What are the procedural steps and pitfalls of the TURBT procedure that should be trained according to an expert panel? Is the Simbla TURBT simulator a feasible and valid simulator for training these steps and avoiding these pitfalls, to complement learning in clinical practice?”

6. What are the technical and non-technical skills necessary for a urologist to show competency in TURBT? Is the newly developed TOCO-TURBT tool a reliable and valid assessment tool?
OUTLINE OF THIS THESIS

In Chapter 2 we investigate the current technical performance of junior and senior residents regarding four basic urological procedures; urethrocystoscopy (UCS), transrectal ultrasound of the prostate (TRUS), transrectal ultrasound-guided prostatic biopsies (TRUSP), and transurethral resection of a bladder tumour (TURBT). We aim to identify concomitant patient safety issues and the specific training needs for these four procedures.

In Chapter 3 we focus on the patient’s point of view, as we study the patient comfort and satisfaction during the diagnostic urological procedures UCS and TRUSP, performed by urologists, junior and senior residents. Furthermore, the degree of interpersonal and communication skills of residents and urologists, as perceived by patients, is investigated. Additionally, the influence of procedure-related factors such as type of scope in UCS and use of anesthetics in TRUSP is evaluated.

In Chapter 4 we aim to gain insight into the current and ideal urological practical skills training. In this chapter we present the outline of the newly designed Dutch Urological Practical Skills (D-UPS) curriculum; a curriculum that provides modular simulation-based training of technical and non-technical basic urological skills in the local hospital setting. Furthermore, design characteristics that could increase the curriculum’s acceptability are assessed.

Chapter 5 evaluates the national implementation of the D-UPS curriculum. The educational impact of the curriculum as perceived by residents and program directors is analyzed, and focus points for improvement of the D-UPS curriculum are provided.

Chapter 6 studies the educational value of the physical ‘Simbla’ TURBT simulator as an educational tool within urological residency training. This chapter describes the results of a TNA that is performed to identify the training objectives for TURBT. Subsequently, the value of the ‘Simbla’ TURBT simulator in achieving these objectives is analysed. Furthermore, face, content and construct validity as well as feasibility and acceptability of the simulator are assessed.

In Chapter 7 we describe the development of a summative assessment tool that measures surgical competency in TURBT. This assessment tool is named the “TOCO-TURBT tool” and is designed by means of a CTA, including clinical observations, hierarchical task decomposition and expert consensus. Subsequently, a validation study was performed in which the TOCO-TURBT tools’ feasibility, content validity, construct validity and reliability was assessed. A total of 51 residents and 25 urologists participated in this validation study.

Finally, the answers to our research questions will be summarized and discussed in Chapter 8. Moreover, this chapter will report our recommendations, practical implications, and suggestions for further research.
REFERENCES


(23) ACGME Program Requirements for Graduate Medical Education in General Surgery. http://www.acgme.org/acgmeweb/Portals/0/ProgramRequirements/440_general_surgery_01012008_07012012.pdf


(54) Dutch Health Care Inspectorate (IGZ. Veilig gebruik van medische technologie krijgt onvoldoende bestuurlijke aandacht in de ziekenhuizen. 2014.


