General discussion and future perspectives
GENERAL DISCUSSION

In this thesis we have aimed to answer two main research questions: ‘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?’

Summary of main finding in light of the literature

In the following three subsections we will discuss our main findings that are presented in the six studies of this thesis, in the light of the literature. This is followed by the answers to our two main research questions in an overall conclusion.

Part 1: Design and content of the D-UPS curriculum.
One of the first steps in the design of a simulation-based training (SBT) curriculum is the identification of training needs and training objectives. In the Dutch Urological Practical Skills (D-UPS) curriculum, the backward design principle of Wiggins and McTighe was applied. This means that the goal and learning outcomes of a training module are determined first, after which learning objectives are established by working backwards from the desired outcomes. In concrete terms this means that a training needs analysis was conducted as the first step in the design of all training modules.

In Chapter 2 this topic was further investigated by assessing the current performance of junior and senior residents regarding basic urological procedures. We showed that this involved a high percentage of unintended events, especially in TRUSP and TURBT, caused by junior and senior residents. Root cause analysis revealed that these events were mainly caused by human factors, in particular verification and skills-based issues. These results can be used in the development of targeted skills training, aiming to optimize quality of care and patient safety.

In Chapter 3 we addressed the patients’ point of view regarding their experiences on comfort and satisfaction during the diagnostic procedures UCS and TRUSP, performed by urologists, junior residents, and senior residents. Patients expressed a high level of satisfaction and comfort during these diagnostic urological procedures, which was not affected by the level of training. Moreover, patients were very satisfied with the degree of communication, provision of information, and after care provided by their physician. These results suggests highly developed ‘interpersonal and communication skills’ for residents in an early stage of urological residency training. This is one of the six core competencies that residency programs need to fulfil according to the requirements of the Accreditation Council for Graduate Medical Education (ACGME). It seems that the
current Dutch curriculum pays sufficient attention to interpersonal and communication skills.

For a successful design of a SBT curriculum it is important to involve the target audience. In Chapter 4 we analyzed the opinions of residents and program directors on current and ideal urological residency training. The results showed that residents are still predominantly trained ‘by doing’, while, ideally, they would train on simulation models first.

Educationalists have described numerous design principles that can be used in the process of designing a urological SBT curriculum. A theoretical framework was provided in the introduction of this thesis, in which the principles of deliberate practice, distributed practice, and whole-task training were highlighted and explained. Zevin and colleagues confirmed that these principles should be included in the design of a comprehensive SBT curriculum.

Moreover, participation should be mandatory. All these aspects were included in the final design of the Dutch Urological Practical Skills training curriculum, presented in Chapter 4 with its main features: 1) training of technical and non-technical basic urological skills; (2) local hospital setting; (3) small groups; (4) use of peer teaching and expert supervision, and (5) yearly recurrence. The acceptability of implementing this SBT curriculum appeared to be high. Design characteristics that would increase its acceptability according to residents and program directors were structured scheduling, the use of peer teaching and high fidelity models.

Part 2: Implementation strategy and educational impact of the D-UPS curriculum

“Change is never easy. You fight to hold on. You fight to let go” - Daniel Stern

Although several surgical SBT curricula have been successfully implemented and certified, the implementation of a new curriculum remains challenging. Establishing acceptability is crucial, and necessitates the early involvement of residents and program directors. The expected difficulties in the implementation of the D-UPS curriculum were described in Chapter 4 of this thesis and included motivation, logistics, and materials. Similar aspects have been previously described in the literature.

In Chapter 5 of this thesis, the D-UPS curriculum was implemented on a national scale and its perceived educational impact was assessed. The motivation of the participants appeared to be high and the vast majority of residents and supervising urologists judged the D-UPS curriculum to be an important addition to current residency training. Unfortunately, the perceived educational impact of training on e.g. knowledge of materials and the ability to anticipate on complications was higher for junior residents than for senior residents. This means that measures are needed that can adjust the level of training to the experience level of residents. To accomplish this, the integration of spiral learning was suggested, which will be further discusses in the
‘future perspectives’ section on page 151. Another focus point for improvement of the D-UPS curriculum after its first implementation was attention to logistics. Structured scheduling is paramount for the successful implementation of a SBT curriculum. The proposed planning for D-UPS curriculum was a weekly 1h training module, similar to the planning of the 4-year curriculum presented by Mc Dougall et al. However, the majority of residents preferred a monthly afternoon of training. Therefore, some hospitals decided to plan one afternoon of training, in which several training modules were clustered. Planning a single hour of training in the busy schedule of residents and urologists incurs the risk that the training will be delayed or even cancelled, leading to a decrease in preparation, participation, and quality of training. This is where theory collides with practical experience, as the principle of ‘distributed practice’ suggests that it would be better to frequently train for shorter periods instead of less frequently for longer periods. Fortunately, the vast majority of participating hospitals succeeded in creating a fixed schedule for the training modules.

Part 3: Design and validation of a high-stakes assessment tool
With the shift from time-based residency training towards competency-based residency training there is a concurrent need for summative assessment tools that enable the measurement of surgical competency. As these tools are not yet available in the field of Urology, there are many questions left unanswered: “When and how to assess, and in what setting?” and “What methodology to use in the development and validation of a high-stakes assessment tool?” In this thesis, we have addressed these questions in relation to the TURBT procedure as the first endoscopic urological procedure for which a high-stakes assessment tool is developed.

In Chapter 6 we focused on the simulation setting. We evaluated the educational value and validity of the physical ‘Simbla’ TURBT simulator as an educational tool to complement training in clinical practice. We showed that the Simbla TURBT simulator is a feasible, acceptable and valid educational tool for training procedural TURBT skills. The Simbla TURBT simulator is the first physical TURBT simulator for which content and construct validity was confirmed. The advantage of training on a physical simulator over a virtual reality simulator is that real instruments and irrigation can be used. This makes the Simbla TURBT simulator suitable to use in high-stakes assessment.

In Chapter 7 of this thesis, we developed a high-stakes assessment tool that measures surgical competency in TURBT: the TOCO-TURBT tool. The TOCO-TURBT tool was designed by means of a cognitive task analysis, including hierarchical task decomposition, clinical observations, and expert consensus. We showed that the TOCO-TURBT tool is a feasible, valid and reliable high-stakes assessment tool to measure surgical competency in TURBT. Two assessors and two procedures appeared to be sufficient to reach a good reliability. We concluded that the TOCO-TURBT tool
has the potential to be used for future certification of skills in TURBT for residents and urologists.

In an attempt to capture all technical and non-technical skills as well as the relevant decision points during a TURBT procedure, a cognitive task analysis (CTA) was conducted. As previously explained, experts have difficulties in explicating the decisions they make during procedures, as their knowledge has become automated. CTA is a valuable method to extract the knowledge of experts and create a gold standard protocol for a procedure. In this, the focus is not only on the overt observable behaviour of the residents, but also on the cognitive functions that underlie the behaviour including non-technical skills. Moreover, we have used expert consensus in the development of the tool, in order to increase its acceptability and content validity.

An important decision in the development of the TOCO-TURBT tool was which type of rating to use. Most well-known are the task specific checklist and the global rating scale, both with their own advantages and disadvantages. The task specific checklist provides focused feedback due to its binary rating (yes/no), but often neglects higher levels of competencies and tends to penalize experts unfairly as they rewards thoroughness instead of clinical competence. The global rating scale on the other hand, is more suitable to capture more holistic components of clinical competence and competencies related to patient safety. Therefore, we chose for a combination of both, making the tool less rigid than task-specific checklists, but still provide specific formative feedback.

The TOCO-TURBT tool was validated in a simulation setting, enabling a true comparison of performance without any interfering confounding factors. However, certain aspects of the procedure and its consequent constituent skills (e.g. bleeding and obtaining haemostasis) could not be simulated and were only assessed in a cognitive way. Therefore, we think that extrapolation of the TOCO-TURBT tool to the clinical setting is desirable in the future, especially for the purpose of certification or recredentialing, as performance in the operating room is the gold standard.

Before the TOCO-TURBT tool can be used as a summative assessment tool, pass/fail scores will have to be set. A pass/fail score is a cut off score along the score scale that defines the boundaries between a competent and non-competent resident. There are several methods that can be used to calculate pass/fail scores, such as the contrasting groups, borderline group, and borderline regression method. Future study should focus on which method is most suitable to obtain reliable and consistent pass/fail scores for the TOCO-TURBT tool.
CONCLUSIONS

‘What is the optimal design, content and implementation strategy for a urological simulation-based training curriculum and what is its educational impact?’

The D-UPS curriculum has incorporated the design principles of deliberate practice, distributed practice, and whole-task training. Characteristics that increase its acceptability are structured scheduling, the use of peer teaching, and high fidelity models. The content of training modules is determined by a training needs analysis including clinical observations.

We carefully conclude that Dutch residents’ interpersonal and communication skills are highly developed in an early stage of training. Nevertheless, present performance of basic urological procedures involves a high percentage of unintended events, especially in TRUSP and TURBT, which are mainly caused by human factors, in particular, verification and skill-based issues. Training modules should focus on these human factors, aiming to reduce the unintended events and optimize patient safety.

The D-UPS curriculum was judged to be an important addition to current residency training, although its perceived educational impact was higher for junior than for senior residents. Thereby, focus points for improvement of the D-UPS curriculum include adaptation of the training level to residents’ level of experience, by the integration of a spiral learning approach, and focus on logistics.

‘How can competency in technical and non-technical skills of TURBT be assessed in a validated, standardized simulation setting?’

The Simbla TURBT simulator has been shown to be a feasible and valid educational tool that can be used in the assessment process. The TOCO-TURBT tool, developed by means of a cognitive task analysis, has been shown to be a feasible, valid, and reliable high-stakes assessment tool to measure surgical competency in TURBT. The use of two assessors and two procedures are sufficient to reach a good reliability of the assessment tool.

Methodological considerations
This thesis has described several studies, all of which contribute to the central theme of how to design, validate and implement simulation based training and assessment. An important strength of this thesis is that we have captured the perspectives of all the stakeholders: residents, program directors and patients. However, a limitation throughout this thesis was the limited numbers of participants that were included in our studies. Concerning the pre-and-post evaluation of the D-UPS curriculum, the limited number of participants was inherent to the fact that the curriculum is implemented in
General discussion and future perspectives

a small country as the Netherlands, with approximately 20-25 residents per training year. Although the number of participants was relatively small, the response rates to our surveys were high and comprised information of all the teaching hospitals where the curriculum was implemented. This ensures a realistic evaluation of the development and implementation of the D-UPS curriculum. Still, for the studies described in chapter 6 and 7, validating a simulator and assessment tool for TURBT, larger numbers of participants are preferable. This is particularly paramount in the future, when pass-fail scores for the TOCO-TURBT tool will be set. In future studies, international collaboration could reach an increased number of participants and make our simulation based training and assessment applicable for an international community.

Throughout this thesis we have used questionnaires and interviews or written qualitative feedback. This might have led to socially desirable answers. To counter this effect, we have anonymized the questionnaires, and the interviews were moderated by an independent educational expert. Furthermore, residents and program directors were interviewed in separate groups to ensure freedom and safety in expressing opinions. Also in Chapter 5, where we assessed the educational value of the D-UPS curriculum, only the self perceived effects were measured by means of a structured quantitative survey. A more objective approach is needed in future validation studies. This is discussed in the next section.

FUTURE PERSPECTIVES

In the next subsections we describe our recommendations regarding future modification of the D-UPS curriculum and its subsequent validation. Moreover, we present our view on the potential use of high-stakes assessment tools when awarding entrustable professional activities (EPAs). Finally, we comment on the use of skills assessment within the selection of new residents and recertification of specialists.

Integration of a spiral learning approach within the D-UPS curriculum

The evaluation of the implementation of a new curriculum leads to new insights and points of improvement that should not be ignored. As Kern and colleagues pointed out:

‘Curriculum development does not usually proceed in sequence, one step at a time. Rather, it is a dynamic, interactive process that continues and the curriculum evolves, based on evaluation results, changes in resources, targeted learners, and the material requiring mastery.’
The most important insight after the first evaluation of the D-UPS curriculum was the discrepancy in perceived educational impact for junior versus senior residents. In the set-up of the D-UPS curriculum, the same training modules were attended by all the residents (junior as well as senior) who worked in a particular teaching hospital. Despite the measures that were taken to raise the training level for senior residents (e.g. use of peer teaching, and possibility to perform modules on a senior level), the educational impact was significantly higher for junior residents compared to senior residents. This reflects the need for measures that adjust the required level of training to the experience level of the particular residents. We suggest a more personalized approach by the integration of spiral learning. In spiral learning, the same skill is taught at each level, but with increasing degrees of complexity and sophistication (fig 1).  

![Spiral Learning Diagram](image)

**Figure 1.** The spiral learning approach

The periodic recycling of the same topics with progressively greater complexity continues exposure to certain topics or skills while facilitating deeper understanding of the subject and promoting more intuitive handling.

To measure progression and assess the need for skills training for experienced, final-year residents or highly technically skilled residents, we suggest the integration of assessment of skills. In this, the concept of ‘entrustable professional activities’ (EPAs) could be valuable. EPAs are defined as “tasks or responsibilities to be entrusted to a resident once sufficient competency is reached to allow for unsupervised practice”. EPA assessment enables supervisors to know when a resident can be trusted to
carry out specific procedures with minimal or no supervision, in different stages of the training. EPA consists of five levels of entrustment that can be linked to the five levels of skills acquisition described in the Dreyfus and Dreyfus model:

Table 1. Stages of skills acquisition (Dreyfus and Dreyfus) linked with levels of entrustment (EPA)

<table>
<thead>
<tr>
<th>Stage of skills acquisition (Dreyfus)</th>
<th>Level of entrustment (EPA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice</td>
<td>1. Observation but no execution, even with direct supervision</td>
</tr>
<tr>
<td>Advanced beginner</td>
<td>2. Execution with direct, proactive supervision</td>
</tr>
<tr>
<td>Competent</td>
<td>3. Execution with reactive supervision, i.e. on request and quickly available</td>
</tr>
<tr>
<td>Proficient</td>
<td>4. Execution with supervision at a distance and/or post hoc</td>
</tr>
<tr>
<td>Expert</td>
<td>5. Supervision provided by the trainee to more junior colleagues</td>
</tr>
</tbody>
</table>

Residents that have demonstrated to master the skills of a certain training module at a certain level of entrustment, could proceed to the next level. Reasonable outcome parameters and accessory competency standards should be defined for each level of expertise/entrustment described by Dreyfus or EPA.

Validation of the D-UPS curriculum

Future validation of the D-UPS curriculum is paramount in the process of innovating educational programs. In this, Kirkpatrick’s training evaluation model could be used. This model distinguishes four levels of evidence: reaction, learning, behaviour, and result (Fig 2.).
In the validation process of the D-UPS curriculum we have now reached level 2 of Kirkpatrick’s training evaluation model. To reach level 3 and 4, it will be necessary to perform an RCT that measures the effect of training on residents’ skills (level 3) and patient outcomes (level 4). It will be challenging to measure the effect of the D-UPS curriculum on patient outcomes, as there are many factors that could potentially influence the end results (e.g., pre and post-operative care, procedure environment, interference of a supervisor). An alternative approach could be to show a relation between SBT and a reduction in number of errors and events in the operating room. For laparoscopic surgery, such an error-rating tool was previously designed and validated.

In the future validation of the D-UPS curriculum, conducting a cost-effectiveness analysis is also of paramount importance. The educational costs have to be weighed up against the cost savings and health care benefits to be able to convince policy makers, universities, hospitals etc. to continue to invest in the D-UPS curriculum. Potential cost savings could be an increase in operating room efficiency and a reduction in operative errors and events. Previous studies have shown significant savings in medical care costs after simulation training, due to a decrease in complication rates.

Use of high-stakes assessment tools when awarding EPAs
At this moment, the Dutch urological curriculum is subject to change. In the renewed curriculum the concept of EPAs will be officially incorporated. The EPA assessments are based on where the resident is on a five point level-of-supervision scale (described on page 155), in which the fourth level is entrustment to effectively perform the EPA without supervision. This raises the question how the level of entrustment is determined. In the literature, entrustment decisions are described as being a complex multifactorial process, and a significant variability was found in how clinicians make entrustment decision. Reported methods are direct observation, discussions with colleagues and even blind faith. In our opinion, a more objective approach would be desirable.

We suggest the use of objective assessment tools, such as the TOCO-TURBT tool, to assess competency standards for each level of entrustment for various procedures. This would lead to an objective evaluation of competency development. To accomplish this, the components of the procedure in question that residents should master in the various entrustment levels should be determined first.

Skills assessment in the selection of new residents
The idea of using skills assessment in the selection of new surgical residents is controversial in the Netherlands. Nevertheless, literature suggests that 5-15% of surgical residents have difficulty in attaining technical competency during their residency training. Moreover, every year a serious number of residents discontinue their surgical training, which results in great expenses for the government. Would this be reason enough to modify the selection process and include skills assessment? We intuitively think that it
might be. However, evidence regarding the predictive value of skills assessment in the selection process and subsequent technical performance is lacking. Therefore, we cannot (yet) recommend the use of skills assessment as a robust selection method for new surgical residents.

Skills assessment in recertification of specialists
What about the assessment of competency on specialist level? Can we assume that specialists (e.g. surgeons, urologists, gynaecologists) maintain their competency throughout the years? Or would their skills deteriorate when they get older? And can we assume that skills are retained after a period of inactivity, for example after a specialists’ sabbatical?

Literature has shown a diminished patient outcome for complex surgery performed by older surgeons. This suggests a need for recertification of skills for older surgeons who practice highly demanding technologies such as (robot-assisted) laparoscopy.

We think that, due to the medical technological developments, the growing emphasis on patient safety, and potential medico-legal consequences for physicians, certification and recertification of specialist will be needed in the near future to ensure skills maintenance and thereby quality of care.

FINAL CONCLUSION

With this thesis we have aimed to make the global developments within surgical education tangible for Dutch Urological residency training. The result is the D-UPS curriculum, which is now implemented on a national scale, and the TOCO-TURBT tool, as the first step towards high-stakes assessment within competency based urological training.

If you want to teach people a new way of thinking, don’t bother trying to teach them. Instead, give them a tool the use of which will lead to new ways of thinking

Richard Buckminster Fuller
REFERENCES


