Exploring the Changing Landscape of Surgical Residency Training

C.J. Hopmans

The studies described in this thesis were performed at the Department of Surgery of the hospitals representing surgical training region IV in the Netherlands, consisting of the Erasmus MC University Medical Center, Rotterdam; Amphia Hospital, Breda; Ikazia Hospital, Rotterdam; Maasstad Hospital, Rotterdam; Reinier de Graaf Groep, Delft; Sint Franciscus Gasthuis, Rotterdam; IJsselland Hospital, Capelle aan den IJssel

Printing of this thesis has been financially supported by:
Erasmus MC
Erasmus MC, Afdeling Heelkunde
Amphia Ziekenhuis

ISBN: 978-90-9030402-1

Cover design and layout: Tania Hopmans - Pearse-Danker
Printing: DPN Rikken Print, Nijmegen, The Netherlands

Exploring the Changing Landscape of Surgical Residency Training

Verkennen van een veranderend chirurgisch opleidingslandschap

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. H.A.P. Pols
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 5 juli 2017 om 09.30 uur

door

Cornelis Johannes Hopmans
geboren op 26-3-1984 te Tholen

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotor:	Prof.dr. J.N.M. IJzermans
Overige leden:	Prof.dr. M. de Hoog Prof.dr. I.H.M. Borel Rinkes Prof.dr. L.P.S. Stassen
Copromotor:	Dr. P.T. den Hoed

TABLE OF CONTENTS

Chapter 1

Introduction, aim and outline of the thesis

Cornelis J. Hopmans

INTRODUCTION

Historically, surgical training has been based on the master-apprenticeship model. In this well-established training model, surgical residents rotated through the various surgical departments and services of generally a single hospital, while gradually taking on increasing levels of autonomy and responsibility under the strict supervision of a role model. This role model, usually the program director, was admired and respected for his attitude and operative skills, and surgical residents clearly identified themselves with his and their hospital's doctrine. The educational principles 'see one, do one, teach one' and 'learning by doing' were the order of the day and a formal and structured training program was often lacking. Assessment of surgical residents' skills took place during informal conversations and was in general based on evidence for operative competence, such as accumulated case numbers and time taken for a surgical procedure.[1,2] Sufficient operative experience was guaranteed by working long hours and participating in as many operations as possible, resulting in an excessive exposure to a variety of surgical procedures. On successful completion of the training program, surgical residents obtained their board certification in general surgery, and due to their broad repertoire of medical-technical skills were adequately prepared for the full scope of tasks being expected to perform in professional practice.

Within the past decade, the structure and format of surgical residency training has changed radically by the introduction of competency-based curricula, the implementation of stringent work hour restrictions and the progressive fragmentation of general surgery into subspecialties. In the next paragraphs, these developments, including their underlying factors, are described in detail. In addition, it is explained how these developments have affected and transformed surgical residency training in the Netherlands.

Introduction of competency-based curricula

The changing needs and expectations of patients and society in the era of internet information,[3] an increased emphasis on debates concerning patient safety,[4] and a rising awareness for accountability, transparency and quality in medicine from both a professional and public point of view,[5] have initiated an international shift toward a more objective and formalized approach of postgraduate medical training.[6-9] To ensure that the next generation of surgeons is well-trained and educated to meet these challenging healthcare demands, and thus guarantee the continued delivery of high-quality patient care, the Association of Surgeons

of the Netherlands, bearing responsibility for the format and content of the training program in general surgery, implemented a systematic, competency-based training and assessment program in 2009, based on the Canadian Medical Education Directives for Specialists (CanMEDS) framework.10 This framework, derived from the Royal College of Physicians and Surgeons of Canada defines 7 competencies, namely Medical Expert, Communicator, Collaborator, Scholar, Health Advocate, Manager and Professional. It has been designed from the perspective that not only medical-technical skills need to be acquired but also a wide range of other general skills must be obtained during postgraduate medical training.[11] To document and evaluate competence progression in an objective, standardized and transparent way, various assessment tools have been adopted on which scores and feedback are assigned, referring to the different CanMEDS competencies.[12-16]

Implementation of work hour restrictions

Within the past decade, work hour restrictions have been enforced by law in most Western countries. In the European Union, implementation of the European Working Time Directive (EWTD) has progressively reduced the average working week from a maximum of 58 hours in 2004, to 56 hours in 2007, to ultimately 48 hours in 2009. The Netherlands, however, received a temporary extension to meet the 48-hour target, with working weeks being restricted to 52 hours in the transition period that lasted until 2011. In addition to this substantial reduction in training time, strict requirements for daily break moments, weekly rest periods and regulation of night and shift-work are stipulated.[17] The aim of the work hour restrictions is to protect the health and safety of healthcare professionals, as well as to reduce the risk to patient safety posed by exhaustion and fatigue resulting from working excessive long hours. Within the medical profession, and in particular surgical specialties, work hour restrictions are considered burdensome because they are believed to have deteriorated the quality of training and education, potentially threatening the high standard of surgical care and safety of patients.[18-20]

Fragmentation of general surgery into subspecialties

Innovations in medical technology and the rapid development of medical knowledge have revolutionized surgical healthcare, but also have expanded the field of general surgery to such an extent that it may be questioned whether a general surgeon will remain able to adequately manage a broad spectrum of surgical conditions and can be assumed competent to perform a wide repertoire of surgical procedures. In addition, an increasingly well-informed public prefers to be treated by surgeons, who are recognized for their expertise on certain fields of disease management.[21] This growing quest for advanced and specialized healthcare is driven by scientific evidence that patient outcomes, particularly for diverse types of cancer

and more complex procedures like aneurysm surgery, are improved if performed by experienced surgeons surrounded by dedicated medical teams in high-volume hospitals.[22,23] Taking into account these developments, in combination with an increased pressure on healthcare budgets, training in general surgery in the Netherlands, often lasting 8 years due to postresidency fellowship training, has been shortened to a 6-year program in which subspecialty training is mandatory and offered during the last 2 years. As a consequence, surgical residents are currently no longer trained as broad-based general surgeons, but as surgical specialists who will primarily be focused on the subspecialty in which training has been received for the rest of their professional careers, such as gastro-intestinal surgery, surgical oncology, trauma surgery or vascular surgery.

AIM AND OUTLINE OF THE THESIS

The aforementioned developments mark the starting point of this thesis in which we aimed to explore how the landscape of surgical residency training has transformed within the past years. The research described in this thesis has been conducted in a training region located in the Southwest of the Netherlands, in which surgical residency training is organized in 1 university hospital (Erasmus MC university Medical Center, Rotterdam) and 6 affiliated district hospitals (Amphia Ziekenhuis, Breda; Ikazia Ziekenhuis, Rotterdam; Maasstad Ziekenhuis, Rotterdam; Reinier de Graaf Groep, Delft; Franciscus Gasthuis & Vlietland, Rotterdam; IJsselland Ziekenhuis, Capelle aan den IJssel).

In **chapter 2** the importance of the various CanMEDS competencies for the competency profile of a surgeon was investigated, using a survey distributed among surgical residents and attending surgeons. In addition, the suitability of the adopted assessment tools to evaluate and monitor competence progression was surveyed.

In **chapter 3** a mixed-methods analysis is presented, which implies that both qualitative and quantitative research methods are used to describe how surgical residents and attending surgeons perceive and deal with the sweeping transformation of surgical training, with its underlying competency framework, assessment tools, work hour restrictions and subspecialty training.

Surgical specialties are unique in a way that not only medical knowledge must be obtained but also operative skills need to be acquired. Within competency-based curricula, the instrument Objective Structured Assessment of Technical Skills (OSATS) is considered the standard for the assessment and evaluation of operative skills.[24] Evidence for the usefulness of this instrument has been demonstrated in experimental study designs,[25,26] but data supporting OSATS application in the operating room are limited.[27-31] In **chapter 4** the validity and reliability of this instrument was examined to evaluate the operative skills of surgical residents, while performing a variety of surgical procedures in an operating theater setting.

Implementation of the EWTD has progressively reduced the working week to an average of 48 hours. In particular surgical specialties have expressed concerns and remain opposed to this significant reduction in training time, primarily as a result of the perception that a decrease in surgical residents' work hours comes at the expense of sufficient operative experience.[32] In **chapter 5**, operating records of the various hospitals within this training region in which all surgical procedures performed by different grades of surgeons are registered, were analyzed for the period 2005-2012. The aim of this study was to answer the question whether implementation of the EWTD has affected the operative experience of surgical residents.

An important aspect of implementing a new training methodology includes monitoring and evaluation. Furthermore, information regarding other important developments, such as the implementation of work hour restrictions and the fragmentation of general surgery into subspecialties, is urgently needed. **Chapter 6** describes the results of a survey that was distributed in both 2011 and 2013. The aim of this successive multicenter survey was to provide an interim analysis on how surgical residents and attending surgeons in the Netherlands perceive the transformation of surgical training.

In **chapter 7** another important aspect related to surgical training was investigated: resident selection. As the aim of surgical training is to adequately prepare the next generation of surgeons to independently perform the full scope of tasks they are expected to carry out in professional practice, selection of surgical residents is a high-stake process entailing the responsibility to correctly identify and recruit successful candidates. However, information on factors that may affect performance during surgical residency training is limited. In this chapter, we aimed to determine whether pre-training variables, derived from structured curriculum vitae, as well as secondary and medical school diplomas, could be associated with indicators for performance during surgical training.

REFERENCES

1. Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. BMJ 2003;327(7422): 1032-1037.
2. Reznick RK. Teaching and testing technical skills. Am J Surg 1993;165(3): 358-361.
3. Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. Med Teach 2007;29(7): 642-647.
4. Frenk J, Chen L, Bhutta ZA, Cohen J, Crisp N, Evans T, et al. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. Lancet 2010;376(9756): 1923-1958.
5. Curry L, Wergin JF. Educating professionals: Responding to new expectations for competence and accountability. Jossey-Bass: San Francisco, CA, 1993; 379.
6. Frank JR, Langer B. Collaboration, communication, management, and advocacy: teaching surgeons new skills through the CanMEDS Project. World J Surg 2003;27(8): 972-978; discussion 978.
7. Brennan MF, Debas HT. Surgical education in the United States: portents for change. Ann Surg 2004;240(4): 565-572.
8. Smith AJ, Aggarwal R, Warren OJ, Paraskeva P. Surgical training and certification in the United Kingdom. World J Surg 2009;33(2): 174-179.
9. Collins JP, Civil ID, Sugrue M, Balogh Z, Chehade MJ. Surgical education and training in Australia and New Zealand. World Journal of Surgery 2008;32(10): 2138-2144.
10. Borel-Rinkes IH, Gouma DJ, Hamming JF. Surgical training in the Netherlands. World J Surg 2008;32(10): 2172-2177.
11. Frank JR. The CanMEDS 2005 Physician Competency Framework. Better standards. Better Physicians. Better care. Ottawa: The Royal College of Physicians and Surgeons of Canada.
12. Lockyer J. Multisource feedback in the assessment of physician competencies. J Contin Educ Health Prof 2003;23(1): 4-12.
13. 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.
14. Norcini JJ, Blank LL, Arnold GK, Kimball HR. The mini-CEX (clinical evaluation exercise): a preliminary investigation. Ann Intern Med 1995;123(10): 795-799.
15. Norman G, Davis D, Painvin A, Lindsay E, Rath D, Ragbeer M. Comprehensive assessment of clinical competence of family/general physicians using multiple measures. Proceedings of the Annual Conference on Research in Medical Education Washington DC: Association of American Medical Colleges 1989: 75-79.

16. Sauve S, Lee HN, Meade MO, Lang JD, Farkouh M, Cook DJ, et al. The Critically Appraised Topic: A Practical Approach to Learning Critical Appraisal. Annals RCPSC 1995;28(7): 396-398.
17. Fitzgerald JE, Caesar BC. The European Working Time Directive: a practical review for surgical trainees. Int J Surg 2012;10(8): 399-403.
18. Axelrod L, Shah DJ, Jena AB. The European Working Time Directive: an uncontrolled experiment in medical care and education. JAMA 2013;309(5): 447-448.
19. Glomsaker TB, Soreide K. Surgical training and working time restriction. Br J Surg 2009;96(4): 329-330.
20. Rosenbaum L, Lamas D. Residents' duty hours--toward an empirical narrative. N Engl J Med 2012;367(21): 2044-2049.
21. Fischer JE. The impending disappearance of the general surgeon. JAMA 2007;298(18): 2191-2193.
22. Birkmeyer JD, Sun Y, Wong SL, Stukel TA. Hospital volume and late survival after cancer surgery. Ann Surg 2007;245(5): 777-783.
23. Holt PJ, Karthikesalingam A, Hofman D, Poloniecki JD, Hinchliffe RJ, Loftus IM, et al. Provider volume and long-term outcome after elective abdominal aortic aneurysm repair. Br J Surg 2012;99(5): 666-672.
24. van Hove PD, Tuijthof GJ, Verdaasdonk EG, Stassen LP, Dankelman J. Objective assessment of technical surgical skills. Br J Surg 2010;97(7): 972-987.
25. Anastakis DJ, Regehr G, Reznick RK, Cusimano M, Murnaghan J, Brown M, et al. Assessment of technical skills transfer from the bench training model to the human model. Am J Surg 1999;177(2): 167-170.
26. Datta V, Bann S, Beard J, Mandalia M, Darzi A. Comparison of bench test evaluations of surgical skill with live operating performance assessments. J Am Coll Surg 2004;199(4): 603-606.
27. 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.
28. Beard JD, Choksy S, Khan S. Assessment of operative competence during carotid endarterectomy. Br J Surg 2007;94(6): 726-730.
29. Hiemstra E, Kolkman W, Wolterbeek R, Trimbos B, Jansen FW. Value of an objective assessment tool in the operating room. Can J Surg 2011;54(2): 116-122.
30. Niitsu H, Hirabayashi N, Yoshimitsu M, Mimura T, Taomoto J, Sugiyama Y, et al. Using the Objective Structured Assessment of Technical Skills (OSATS) global rating scale to evaluate the skills of surgical trainees in the operating room. Surg Today 2013;43(3): 271-275.
31. 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.
32. Kairys JC, McGuire K, Crawford AG, Yeo CJ. Cumulative operative experience is decreasing during general surgery residency: a worrisome trend for surgical trainees? J Am Coll Surg 2008;206(5): 804-811; discussion 811-803.

Chapter 2

Surgeons' attitude toward a competency-based training and assessment program: Results of a multicenter survey

Cornelis J. Hopmans
Pieter T. den Hoed
Iris Wallenburg
Lijckle van der Laan
Erwin van der Harst
Maarten van der Elst
Guido H.H. Mannaerts
Imro Dawson
Jan J.B. van Lanschot
Jan N.M. IJzermans

Adapted from: J Surg Educ. 2013 Sep-Oct;70(5):647-54

ABSTRACT

Background: Surgical training programs are increasingly focused on the development and evaluation of professional competencies. Also in the Netherlands, competency-based curricula have been introduced to restructure postgraduate medical training. The current surgical residency program is based on the Canadian Medical Education Directives for Specialists (CanMEDS) competencies. To evaluate competence progression, assessment tools have been incorporated. This study examined the attitude of surgical residents and attending surgeons toward a competency-based training and assessment program used to restructure training in general surgery in the Netherlands in 2009.

Methods: In 2011, all surgical residents (n=51) and attending surgeons (n=108) in a training region, consisting of 7 hospitals, were surveyed. Participants were asked to rate the importance of the CanMEDS competencies and the suitability of the adopted assessment tools. Items were rated on a 5-point Likert scale and considered relevant when ≥80% of the respondents rated an item with a score of 4 or 5 (indicating a positive attitude). Reliability was evaluated by calculating Cronbach's α coefficient. The Mann-Whitney test was applied to assess differences between groups.

Results: The response rate was 88% (n=140). The CanMEDS framework demonstrated good reliability (Cronbach's α=0.87). However, the importance of the competencies Manager (78%) and Health Advocate (70%) was undervalued. The assessment tools failed to achieve an acceptable reliability (Cronbach's α=0.55) and were predominantly considered unsuitable for evaluation of the CanMEDS competencies. Exceptions were the assessment tools In-Training Evaluation Report (91%) and Objective Structured Assessment of Technical Skills (82%). No significant differences were found between the surgical residents and the attending surgeons.

Conclusion: Two years after reforming the training program in general surgery in the Netherlands, surgical residents and attending surgeons did not recognize the importance of all CanMEDS competencies and consider the assessment tools generally unsuitable for competence evaluation purposes.

INTRODUCTION

Over the past decade, in many Western countries, including the United States,[1] the United Kingdom,[2] and Canada[3] the format and structure of surgical training has been reformed. The changing needs and expectations of patients and society in the era of internet information,[4] evolving medical technology,[5] the implementation of stringent work hour restrictions,[6] alterations in the composition of the medical workforce,[7] as well as the renewed balance of professional ambition weighed against lifestyle issues and private responsibilities[8] are all developments driving the reform of postgraduate medical training. Against the background of these developments, it remains the responsibility of professional surgical associations to ensure that the next generation of surgeons is well-trained and educated to meet challenging future healthcare demands, and thus guarantee the continued delivery of high-quality patient care.

In 2009, the Association of Surgeons of the Netherlands implemented a competency-based curriculum based on the Canadian Medical Education Directives for Specialists (CanMEDS) framework.[9] Also other countries, including Denmark,[10] Finland,[11] Germany,[12] Australia, and New Zealand,[13] have embraced the CanMEDS framework to restructure their surgical training programs. The framework, derived from the Royal College of Physicians and Surgeons of Canada, defines 7 roles. It has been designed from the perspective that not only medical-technical skills need to be acquired but also a wide range of other general skills, related to the CanMEDS roles, must be obtained during postgraduate medical training. An overview and definition of these roles, in the Netherlands called competencies, is shown in Table 1.[14]

Having defined a competency framework, surgical residents are challenged to develop and demonstrate achievement of these competencies.[15] To document and evaluate this in a transparent and objective way, various assessment tools have been adopted.[16-20] These assessment tools, collected in a portfolio, are structured checklists with room for feedback on which assessors have to assign scores, referring to the different CanMEDS competencies. As a result, assessment tools can confirm whether a certain level of competence has been achieved and whether the performance shown is in accordance with the predefined training requirements. Throughout the entire training, these instruments have to be used a minimum number of times, in a variety of clinical situations, according to a specific assessment schedule.

Table 1. Overview and definition of the Canadian Medical Education Directives for Specialists (CanMEDS) competency framework.

CanMEDS competencies	Definition
Medical Expert	As *Medical Experts*, physicians integrate all of the CanMEDS roles, applying medical knowledge, clinical skills, and professional attitudes in their provision of patient-centered care.
Communicator	As *Communicators,* physicians effectively facilitate the doctor-patient relationship and the dynamic exchanges that occur before, during, and after the medical encounter.
Collaborator	As *Collaborators*, physicians effectively work within a healthcare team to achieve optimal patient care.
Scholar	As *Scholars*, physicians demonstrate a lifelong commitment to reflective learning, as well as the creation, dissemination, application and translation of medical knowledge.
Health Advocate	As *Health Advocates*, physicians responsibly use their expertise and influence to advance the health and well-being of individual patients, communities, and populations.
Manager	As *Managers*, physicians are integral participants in healthcare organizations, organizing sustainable practices, making decisions about allocating resources, and contributing to the effectiveness of the healthcare system.
Professional	As *Professionals*, physicians are committed to the health and well-being of individuals and society through ethical practice, profession-led regulation, and high personal standards of behaviour.

This description is adapted from Frank JR. (Ed). The CanMEDS 2005 Physician Competency Framework. Better standards. Better Physicians. Better care. Ottawa: The Royal College of Physicians and Surgeons of Canada.[14]

Successful implementation of a new training methodology is highly dependent on the support of its users. Therefore, it is essential that soon after the implementation feedback is obtained to ensure a proper evaluation. In addition, little is known on the perceived importance of the CanMEDS competencies and suitability of assessment tools within different medical specialties. From this perspective, we conducted a multicenter survey among surgical residents and attending surgeons in a training region, located in the Southwest of the Netherlands. The aim was to examine the attitude toward a competency-based training and assessment program used to restructure training in general surgery.

METHODS

Context of the study

In the Netherlands, training in general surgery is organized in 8 regions, each consisting of 1 university hospital and several affiliated district hospitals. To become a board certified surgeon, a 6-year training program has to be completed, of which 2 years are spent in a university hospital and 4 years in a district hospital. During the first 4 years, training in general surgery is offered. The remaining 2 years consist of dedicated training in one of the following subspecialties: gastrointestinal surgery, surgical oncology, trauma surgery or vascular surgery. In addition, subspecialty training is offered in focus areas such as lung surgery and pediatric surgery. As a result, surgical residents are no longer trained to become broad-based general surgeons, but surgical specialists who will primarily be focused on the subspecialty in which they received training for the rest of their professional careers.

Study design

This study was conducted in a training region in which surgical residency training is covered by 1 university hospital and 6 district hospitals. In November 2011, a survey was distributed electronically (www.surveymonkey.com) and by regular post to all surgical residents (n=51) and attending surgeons (n=108). An introductory letter was enclosed with the survey, explaining the purpose of the study and assuring anonymity and confidentiality. Participation was voluntary. Participants were invited to respond within 1 month. An e-mail reminder was sent if surveys were not returned within 2 weeks after receipt.

Survey

The survey contained 2 sections. In the first section, participants were asked to indicate the importance of the CanMEDS competencies (Medical Expert, Communicator, Collaborator, Manager, Health Advocate, Scholar and Professional) for the competency profile of a surgeon. Scores were rated on a 5-point Likert scale (1= unimportant and 5= very important). In the second section, participants were presented with the assessment tools, including In-Training Evaluation Report (ITER), 360-degree assessment, Critically appraised topic, Case-based discussion, mini-Clinical evaluation exercise, Assessment of an operation report, Report of a complicated hospitalization, and Objective Structured Assessment of Technical Skills (OSATS).16-20 A brief description of these tools is provided in Table 2, showing which

competencies can be evaluated using the different assessment tools.[21] Participants were asked to indicate the suitability of these instruments to evaluate and monitor competence progression. Again, scores had to be rated on a 5-point Likert scale (1= extremely unsuitable and 5= very suitable). Finally, background questions were asked about the participants' gender, age, seniority and type of hospital employment. All returned surveys were analyzed.

Statistical analysis

Statistical analyses were performed with SPSS version 17.0 (SPSS Inc., Chicago, Illinois, US). Descriptive statistics were calculated for all survey items, including mean values (standard deviations [SD]), frequencies and ranges. To establish the reliability of the CanMEDS framework and the assessment tools, the internal consistency reliability (Cronbach's α) was calculated. A Cronbach's $\alpha \geq 0{,}70$ was considered an indication of acceptable reliability.[22] The Cronbach's α if item deleted statistics were checked to determine whether the overall value of α would increase if CanMEDS competencies or assessment tools were removed. If so, this would indicate that these items cause a decrease in the value of α, thus affecting overall reliability, and therefore should be deleted. To assess whether survey items were considered relevant, the content validity index was calculated. This index was defined as the proportion of respondents rating a survey item with a score of 4 or 5 (indicating a positive attitude) on the 5-point Likert scale. Since there is no agreement about what content validity index is acceptable, in accordance with expert guidelines, a cut-off value of 80% was chosen.[23,24] The Mann-Whitney test was applied to analyze differences in the assigned sores between different groups of respondents. The Bonferroni adjustment was used to counteract the problem of multiple testing, which specified the significance level at 0.003.

Table 2. Description of the assessment tools introduced by the Association of Surgeons of the Netherlands to evaluate the CanMEDS competencies.[21]

Assessment Tools	Definition	CanMEDS competencies						
		ME	COM	COL	SCH	HA	MAN	PRO
In-training evaluation report	Periodic review of the portfolio for completeness and evaluation of competence progression	X	X	X	X	X	X	X
360-degree assessment	Evaluation of daily functioning by colleagues, nursing staff and other healthcare workers		X	X				X
Critically appraised topic	Critical appraisal of a scientific article relating to a focused clinical question derived from a real-patient encounter				X			
Case-based discussion	Evaluation of clinical reasoning and decision-making skills after the presentation of a managed clinical case	X			X			
mini-Clinical evaluation exercise	Observation and assessment of clinical skills, attitudes and behaviour during patient encounters, clinical handovers or multidisciplinary meetings	X	X	X		X		X
Assessment of an operation report	Critical assessment of an operation report	X					X	X
Report of a complicated hospitalization	Review of a patient record to evaluate medical record keeping and correspondence	X					X	X
Objective structured assessment of technical skills	Assessment of surgical residents' operative skills in the operating theater.	X						

ME (Medical Expert); COM (Communicator); COL (Collaborator); SCH (Scholar); HA (Health Advocate); MAN (Manager); PRO (Professional)

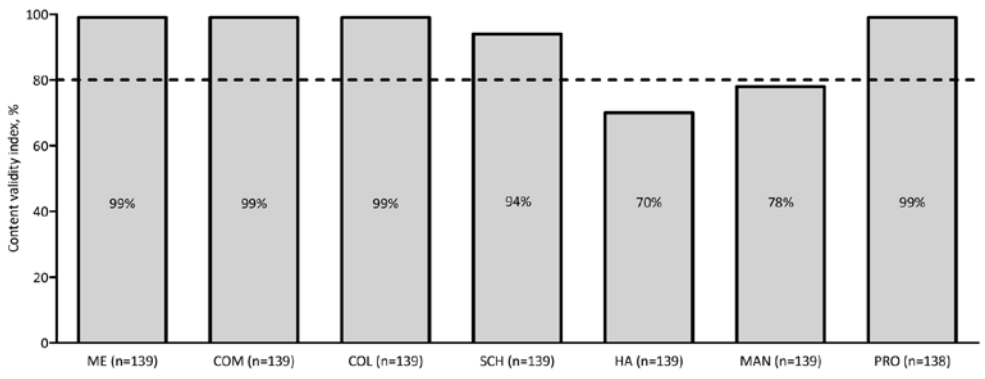
RESULTS

In total, 140 out of 159 invited participants (88%) responded to the survey, including 43/51 (84%) surgical residents and 97/108 (90%) attending surgeons. Missing value analysis revealed that 129 respondents completed all survey items. For the 43 responding surgical residents, the mean age was 31.6 years (range 27-37 years), 26 were male and 17 were female. Forty-three surgical residents (79%) were employed in district hospitals. The 97 responding attending surgeons had a mean age of 44.9 years (range 33-63 years), 84 were male and 13 were female. Thirty-three attending surgeons (34%) worked in the university hospital.

Importance of the CanMEDS competencies

The content validity indices for the 7 competencies of the CanMEDS framework are shown in Figure 1. Overall, not all competencies were considered important for the competency profile of a surgeon. The respondents recognized the importance of the competencies Medical Expert (99%), Communicator (99%), Collaborator (99%), Scholar (94%), and

Figure 1. Content validity indices for the CanMEDS competencies: ME (Medical Expert); COM (Communicator); COL (Collaborator); SCH (Scholar); HA (Health Advocate); MAN (Manager); PRO (Professional).



Content validity was defined as the proportion of respondents rating a survey item with a score of 4 or 5 (on a 5-point Likert scale). Competencies were considered important for the competency profile of a surgeon when the content validity index was ≥80%, indicated by the interrupted line.

Professional (99%). However, the importance of the competencies Manager (78%) and Health Advocate (70%) was undervalued. Reliability analysis of the CanMEDS framework revealed a Cronbach's α of 0.87. The Cronbach's α if item deleted statistics for individual CanMEDS competencies ranged from 0.84-0.87, indicating that the framework had a high level of internal consistency and individual competencies were valued as a coherent construct. As shown in Table 3, no significant differences were found in the assigned ratings of importance for any of the CanMEDS competencies between the group of surgical residents and the group of attending surgeons. There was consensus between both groups, with the competency Medical Expert considered the most important (surgical residents: mean=4.74, SD=0.45; attending surgeons: mean=4.78, SD=0.55) and the competency Health Advocate valued least important (surgical residents: mean=3.83, SD=0.73; attending surgeons: mean=3.74, SD=0.68). Data were further analyzed for gender and hospital type, but no significant differences between respondents of these different subgroups were found.

Suitability of the assessment tools

In the second part of the survey, it was asked to indicate the suitability for each of the assessment tools to evaluate and monitor progression in achieving the CanMEDS competencies. Only the instruments ITER and OSATS were rated suitable, with content validity indices of 91% and 82%, respectively. The other assessment tools, including 360-degree assessment (60%), Critically appraised topic (55%), Case-based discussion (64%), mini-Clinical evalu-

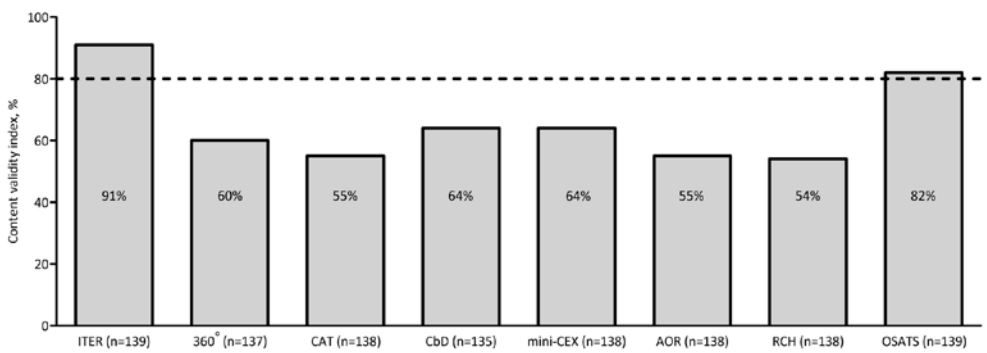
Table 3. Perceived importance of the CanMEDS competencies for the competency profile of a surgeon as scored by the group of surgical residents and attending surgeons.

	Surgical Residents			Attending Surgeons			
CanMEDS competencies	**N**	**mean**	**SD**	**N**	**mean**	**SD**	**p-value**
Medical Expert	42	4.74	0.45	97	4.78	0.55	0.329
Communicator	42	4.67	0.48	97	4.53	0.61	0.215
Collaborator	42	4.55	0.50	97	4.51	0.61	0.861
Scholar	42	4.29	0.60	97	4.27	0.64	0.968
Health Advocate	42	3.83	0.73	97	3.74	0.68	0.580
Manager	42	4.02	0.72	97	3.93	0.71	0.564
Professional	42	4.45	0.55	96	4.50	0.63	0.486

Data are presented as means and standard deviations (SD)

ation exercise (64%), Assessment of an operation report (55%), and Report of a complicated hospitalization (54%) failed to achieve content validity indices ≥80% and were consequently considered unsuitable for competence evaluation. These results are presented in Figure 2. The Cronbach's α coefficient for the 8 assessment tools was 0.55, indicating that the internal consistency reliability of the assessment tools was considered poor. When Cronbach's α if item deleted statistics were taken into account, the overall α did not substantially change (range 0.45-0.58), meaning that the assessment tools were generally perceived unreliable for the evaluation of competence progression. No significant differences were present in the assigned suitability scores between the group of surgical residents and the group of attending surgeons. The instruments ITER (surgical residents: mean=4.28, SD=0.55; attending surgeons: mean=4.07, SD=0.57) and OSATS (surgical residents: mean=4.06, SD=0.72; attending surgeons: mean=3.99, SD=0.73) received the highest suitability scores from both groups (Table 4). No significant differences between respondents of different subgroups were identified.

Figure 2. Content validity indices for the assessment tools: ITER (In-training evaluation report); 360° (360-degree evaluation); CAT (critically appraised topic); CbD (case-based discussion), mini-CEX (mini-clinical evaluation exercise); AOR (assessment of an operation report); RCH (report of a complicated hospitalization); OSATS (objective structured assessment of technical skills).



Content validity was defined as the proportion of respondents rating a survey item with a score of 4 or 5 (on a 5-point Likert scale). Assessment tools were considered suitable to assess surgical residents' competence progression when the content validity index was ≥80%, indicated by the interrupted line.

Table 4. Perceived suitability of the assessment tools to evaluate competence progression as scored by the group of surgical residents and attending surgeons.

	Surgical Residents			Attending Surgeons			
Assessment Tools	**N**	**mean**	**SD**	**N**	**mean**	**SD**	**p-value**
ITER	43	4.28	0.55	96	4.07	0.57	0.049
360°	43	3.56	0.91	94	3.70	0.85	0.326
CAT	43	3.30	0.74	95	3.59	0.64	0.049
CbD	41	3.32	0.79	94	3.70	0.65	0.006
mini-CEX	43	3.42	0.88	95	3.66	0.75	0.089
AOR	43	3.42	0.93	95	3.45	0.78	0.932
RCH	43	3.23	0.90	95	3.63	0.76	0.023
OSATS	43	4.06	0.72	96	3.99	0.73	0.453

Data are presented as means and standard deviations (SD).

ITER (In-training evaluation report); 360° (360-degree evaluation); CAT (critically appraised topic); CbD (case-based discussion), mini-CEX (mini-clinical evaluation exercise); AOR (assessment of an operation report); RCH (report of a complicated hospitalization); OSATS (objective structured assessment of technical skills)

DISCUSSION

This study describes the attitude of surgeons toward a competency-based curriculum that has been implemented to reform training in general surgery in the Netherlands. Like the postgraduate medical training programs of several European[10-12] as well as non-European countries,[13] the surgical residency program in the Netherlands is based on the competencies of the CanMEDS framework. To document and evaluate competence progression, various assessment tools have been adopted.[16-21] By conducting a multicenter survey, we examined how surgical residents and attending surgeons perceived the importance of the CanMEDS competencies and the suitability of the assessment tools.

The results of this study showed a high level of internal consistency (Cronbach's α=0.87) for the CanMEDS framework. However, not all competencies were considered important for the competency profile of a surgeon, undervaluing the importance of the competencies Manager (78%) and Health Advocate (70%). In part, it might be due to the recent implementation of the competency-based training program that not all respondents were well-aware of the meaning of all CanMEDS competencies. Also, results might have been different if the competencies were accompanied by detailed descriptions explaining their underlying aspects. Another explanation for this finding has been suggested by Ringsted et al.[25] In their survey among Danish physicians, differences were found in the assigned ratings of importance for several of the CanMEDS competencies between 4 different kind of specialty groups (laboratory specialties; technical specialties; cognitive specialties; general and social medicine), highlighting the possibility that different competency profiles could be relevant for different kinds of medical specialties.

No significant differences were observed between the group of surgical residents and the group of attending surgeons regarding the perceived importance for any of the CanMEDS competencies. In contrast to these findings, a UK survey reported a discrepancy in the attributed levels of importance to the CanMEDS competencies between surgeons based on seniority. In that particular study, surgical trainees attributed lower importance to the competencies Manager, Communicator, Collaborator, and Professional compared to consultant surgeons.[26] In another study from the UK, conducted among consultant surgeons in Scotland, the importance of general skills, comparable to the roles of CanMEDS framework, was valued more highly in surgical trainees than medical-technical skills.[27] In contrast, in this study the competency Medical Expert was valued as most the important, supporting the traditional view by surgeons that it is paramount to have adequate medical knowledge and excellent surgical skills.

The competency Health Advocate was scored as least important. This result is consistent with previous studies, which also showed lowest ratings for this competency.[25,28] Moreover, it might be of concern that other studies examining the importance of the CanMEDS competencies have reported that surgical residents perceive competencies, in addition to Medical Expert, less adequately taught.[29-31] Perhaps of even greater concern is that medical faculty felt uncertainty about their role in teaching and assessing general competencies.[32] This raises the question whether the current medical specialists are adequately skilled to teach and assess competencies they never learned during their own training. Therefore, we propose that educational resources need to be invested, especially for competencies that are currently undervalued in postgraduate medical training programs. For example, mandatory courses in

communication skills and management have already been implemented in Denmark and the UK, respectively.[10,33]

In addition to the CanMEDS framework, a variety of assessment tools was adopted to document and evaluate competence development in daily clinical activities performed by surgical residents. As shown in Table 2, most assessment tools are mainly focused on the evaluation of the competency Medical Expert, while the number of instruments focusing on the evaluation of the other competencies is relatively limited. This was also observed in a survey among Canadian program directors, showing that assessment of the competency Medical Expert was perceived satisfied, but that there was dissatisfaction and concern about assessment of the competencies, in addition to Medical Expert.[34]

Another important finding of this study is that the assessment tools predominantly were considered unsuitable for competence assessment. As with the CanMEDS framework, an explanation for this finding could be the recent implementation of the competency-based curriculum or unfamiliarity with the use of assessment tools. Of all assessment tools, only 2 achieved a content validity index ≥80%, namely ITER and OSATS. The instrument ITER (91%) is a periodic meeting with the program director who reviews the content of a surgical resident's portfolio for completeness, allowing to distract whether competency goals over a certain period have been achieved. Also in Canada, program directors heavily rely on this instrument to evaluate competence progression.[35] The instrument OSATS (82%) is used for the evaluation of surgical skills. Not surprisingly, there are great similarities between these 2 instruments and the way surgical residents were traditionally assessed, namely during informal conversations and based on accumulated case numbers. This confirms that personal interactions with attending surgeons and evidence for surgical skills are still considered the foundation for surgical residency training by the respondents.

Regarding the assessment tools that were valued unsuitable to evaluate competence progression, it is particularly worrying that these instruments were adopted to evaluate competencies other than that of Medical Expert. Should therefore more assessment tools have to be adopted, or be invested in the development of assessment tools that may possibly be able to evaluate achievement of general competencies? First, we should ask ourselves whether competence achievement, especially of competencies in addition to Medical Expert, can be measured with assessment tools. In a systematic review, it was concluded that the current literature provides no evidence that assessment tools can evaluate competencies independently of one another. Additionally, it was recommended to use competencies to direct and coordinate feedback rather than struggle to develop assessment tools to measure competen-

cies directly.[36] We share this view and add that the emphasis of competency-based training should be focused on the development of professional competencies instead of finding ways to evaluate them. Moreover, we believe that it is hard to assign the presence or absence of competencies based on circling scores on an assessment tool.

Possible limitations of this study are the small study population and the fact that this study was conducted in a single surgical training region of the Netherlands. Furthermore, a disadvantage of survey methodology is that solely quantitative information is provided, whereas a qualitative study design could provide more in-depth information and underlying thoughts of respondents. However, in view of the high response rate (88%) and the multicenter approach, we have no reason to believe that in other training regions of the Netherlands the CanMEDS competencies or the adopted assessment tools would be valued differently. To investigate whether the respondents' attitude toward the CanMEDS framework and the assessment tools has changed over time, a future survey would need to be conducted.

CONCLUSION

This survey has provided important feedback from surgical residents and attending surgeons in a training region of the Netherlands on the training methodology currently used to structure training in general surgery. Two years after the implementation of a competency-based training and assessment program, surgical residents and attending surgeons did not recognize that all CanMEDS competencies are equally important for the competency profile of a surgeon, undervaluing the competencies Manager and Health Advocate. The adopted assessment tools were predominantly considered unsuitable to measure competence progression, suggesting that adequate evaluation of the CanMEDS competencies will remain a future challenge. Briefly, these findings underline that the implementation of a competency-based training and assessment program is a challenging task. To ensure that the next generation of medical specialists will receive tailored high-quality training and education, continuous evaluation of postgraduate medical training programs is essential.

REFERENCES

1. Brennan MF, Debas HT. Surgical education in the United States: portents for change. Ann Surg. 2004 Oct;240(4):565-72.
2. Smith AJ, Aggarwal R, Warren OJ, Paraskeva P. Surgical training and certification in the United Kingdom. World J Surg. 2009 Feb;33(2):174-9.
3. Frank JR, Langer B. Collaboration, communication, management, and advocacy: teaching surgeons new skills through the CanMEDS Project. World J Surg. 2003 Aug;27(8):972-8; discussion 8.
4. Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. Med Teach. 2007 Sep;29(7):642-7.
5. Jones R, Higgs R, de Angelis C, Prideaux D. Changing face of medical curricula. Lancet. 2001 Mar 3;357(9257):699-703.
6. Doctors' training and the European Working Time Directive. Lancet. 2010 Jun 19;375(9732):2121.
7. Phillips SP, Austin EB. The feminization of medicine and population health. JAMA. 2009 Feb 25;301(8):863-4.
8. Davenport DL, Henderson WG, Hogan S, Mentzer RM, Jr., Zwischenberger JB, Participants in the Working Conditions of Surgery R, et al. Surgery resident working conditions and job satisfaction. Surgery. 2008 Aug;144(2):332-8 e5.
9. Borel-Rinkes IH, Gouma DJ, Hamming JF. Surgical training in the Netherlands. World J Surg. 2008 Oct;32(10):2172-7.
10. Mortensen L, Malling B, Ringsted C, Rubak S. What is the impact of a national postgraduate medical specialist education reform on the daily clinical training 3.5 years after implementation? A questionnaire survey. BMC Med Educ. 2010;10:46.
11. Litmanen T, Ruskoaho J, Vanska J, Halila H, Patja K. Does the need for professional competencies change during the physician's career?--A Finnish national survey. Med Teach. 2011;33(5):e275-80.
12. Kadmon M, Busemann A, Euteneier A, Gawad K, Grone J, Berberat P. [Modular postgraduate training in surgery - a national concept with future]. Zentralbl Chir. 2012 Apr;137(2):138-43.
13. Collins JP, Civil ID, Sugrue M, Balogh Z, Chehade MJ. Surgical education and training in Australia and New Zealand. World Journal of Surgery. 2008 Oct;32(10):2138-44.
14. Frank JR. The CanMEDS 2005 Physician Competency Framework. Better standards. Better Physicians. Better care. Ottawa: The Royal College of Physicians and Surgeons of Canada.
15. Epstein RM. Assessment in medical education. N Engl J Med. 2007 Jan 25;356(4):387-96.
16. Lockyer J. Multisource feedback in the assessment of physician competencies. J Contin Educ Health Prof. 2003 Winter;23(1):4-12.

17. 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 Feb;84(2):273-8.
18. Norcini JJ, Blank LL, Arnold GK, Kimball HR. The mini-CEX (clinical evaluation exercise): a preliminary investigation. Ann Intern Med. 1995 Nov 15;123(10):795-9.
19. Norman G, Davis D, Painvin A, Lindsay E, Rath D, Ragbeer M. Comprehensive assessment of clinical competence of family/general physicians using multiple measures. Proceedings of the Annual Conference on Research in Medical Education Washington DC: Association of American Medical Colleges. 1989:75-9.
20. Sauve S, Lee HN, Meade MO, Lang JD, Farkouh M, Cook DJ, et al. The Critically Appraised Topic: A Practical Approach to Learning Critical Appraisal. Annals RCPSC. 1995;28(7):396-8.
21. Association of Surgeons of the Netherlands. Structuur Curriculum Heelkunde voor Reflectieve Professionals (SCHERP, version 2009). General Surgical Residency Training Program in the Netherlands. Available from: www.heelkunde.nl.
22. Brinkman WP. Design of a Questionnaire Instrument. Handbook of Mobile Technology Research Methods: Nova Publisher; 2009. p. 31-57.
23. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health. 2006 Oct;29(5):489-97.
24. Grant JS, Davis LL. Selection and use of content experts for instrument development. Res Nurs Health. 1997 Jun;20(3):269-74.
25. Ringsted C, Hansen TL, Davis D, Scherpbier A. Are some of the challenging aspects of the CanMEDS roles valid outside Canada? Med Educ. 2006 Aug;40(8):807-15.
26. Arora S, Sevdalis N, Suliman I, Athanasiou T, Kneebone R, Darzi A. What makes a competent surgeon?: experts' and trainees' perceptions of the roles of a surgeon. Am J Surg. 2009 Nov;198(5):726-32.
27. Baldwin PJ, Paisley AM, Brown SP. Consultant surgeons' opinion of the skills required of basic surgical trainees. Br J Surg. 1999 Aug;86(8):1078-82.
28. Ortwein H, Knigge M, Rehberg B, Hein OV, Spies C. Validation of core competencies during residency training in anaesthesiology. Ger Med Sci. 2011;9:Doc23.
29. Card SE, Snell L, O'Brien B. Are Canadian General Internal Medicine training program graduates well prepared for their future careers? BMC Med Educ. 2006;6:56.
30. Freiman A, Natsheh A, Barankin B, Shear NH. Dermatology postgraduate training in Canada: CanMEDS competencies. Dermatol Online J. 2006;12(1):6.
31. Kholdebarin R, Helewa RM, Hochman DJ. Evaluation of a regional acute care surgery service by residents in general surgery. J Surg Educ. 2011 Jul-Aug;68(4):290-3.
32. Verma S, Flynn L, Seguin R. Faculty's and residents' perceptions of teaching and evaluating the role of health advocate: a study at one Canadian university. Acad Med. 2005 Jan;80(1):103-8.

33. Day M. Health service management: The rise of the doctor-manager. BMJ. 2007 Aug 4;335(7613):230-1.
34. Chou S, Cole G, McLaughlin K, Lockyer J. CanMEDS evaluation in Canadian postgraduate training programmes: tools used and programme director satisfaction. Med Educ. 2008 Sep;42(9):879-86.
35. Chou S, Lockyer J, Cole G, McLaughlin K. Assessing postgraduate trainees in Canada: are we achieving diversity in methods? Med Teach. 2009 Feb;31(2):e58-63.
36. Lurie SJ, Mooney CJ, Lyness JM. Measurement of the general competencies of the accreditation council for graduate medical education: a systematic review. Acad Med. 2009 Mar;84(3):301-9.

Chapter 3

Repairing reforms and transforming professional practices: A mixed-methods analysis of surgical training reform

Iris Wallenburg
Cornelis J. Hopmans
Martina Buljac-Samardzic
Pieter T. den Hoed
Jan N.M. IJzermans

Adapted from: Journal of Professions and Organization. 2016 Mar;3(1):86-102

ABSTRACT

Information on how medical professionals give shape to external policy reforms and contemporary transitions in postgraduate medical training is limited. Using a mixed-methods design, the reform of surgical residency training in Netherlands was examined. Data were derived from semi-structured interviews (n=26), observations during meetings of an expert group (n=7), and a questionnaire distributed among surgical residents and attending surgeons, yielding a response rate of 88% (140/159). Informed by the sociological literature on medical education and changing professionalism, this study revealed that both surgical residents and attending surgeons are ambivalent toward the reform of surgical residency training with its underlying competency framework, assessment tools, subspecialty training and work hour restrictions. However, in daily clinical practice 'repair work' is conducted in order to respond to and deal with the sweeping transformation, as well as to soften unwanted effects. This goes beyond restoring the status quo in the sense of re-establishing the old, instead, it is an interactive process; implementing modern training methodologies and adapting to external policy reforms while simultaneously seeking convenient and feasible solutions to preserve certain values and related practices.

INTRODUCTION

In the past decade, the changing nature of professional practices has been the subject of debate among scholars working in the field of sociology. Various authors have examined the reshaped relationships and boundaries between institutional key players, such as government, health policy makers and professional associations. The growing body of literature on 'organized professionalism' highlights the mediation and hybridization between organizations and professions in daily practice (Muzio and Kirkpatrick 2011; Noordegraaf 2011; Postma, Oldenhof and Putters 2015). New patterns of professionalism are depicted as professions transform from self-governing entities into skilled and intelligible service providers in professional organizations, acting within and shaped through organizational logic (Faulconbridge and Muzio 2008; Muzio and Kirkpatrick 2011; Kipping and Kirkpatrick 2013). On the one hand, it is argued that this transformation has resulted in the loss of autonomy for individual practitioners as their work becomes more circumscribed by clinical audits, protocols and guidelines (Levay and Waks 2009; Evetts 2011). Others, conversely, argue that professionals adapt to reforms to serve their own interests (Levay and Waks 2009; Waring and Currie 2009) or to stress the ability the profession to restore disrupted institutional arrangements and re-establish the status quo (Micelotta and Washington 2013).

The aim of this study is to outline how professional–organizational relationships have evolved and are actively given shape in everyday practice by examining a rather under-researched topic: the reform of postgraduate medical training. This topic is critical to study (changing) professionalism as it is generally seen as way of institutional reproduction: transferring knowledge and skills, as well as reinforcing legitimacy, while reproducing advantageous institutional arrangements (Micelotta and Washington 2013). Postgraduate medical training facilitates entrance to the medical profession and is aimed to socialize young physicians in the ethical values, cultural norms and status of the profession. Historically, postgraduate medical training is characterized by reliance on expert knowledge gained through lengthy periods of dedicated hands on training and socialization (Bosk 1979; Hafferty 2000). Contemporary reforms, which stretch out across national boundaries and institutional contexts (Ringsted et al. 2006), include a shift from the traditional master-apprenticeship training model based on the educational principles of learning-by-doing and role modeling, toward structured, competency-based curricula based on modern educational insights and standardized performance assessments (Ludmerer and Johns 2005; Bolton, Muzio and Boyd-Quinn 2011; Wallenburg, Pols and de Bont 2015). The competency model encapsulates both medical-technical skills

and general competencies, such as communication and collaboration, with the aim to better equip medical doctors for a changing healthcare context. Hence, the reform of postgraduate medical training is part of a wider trend toward greater involvement of medical doctors in organizational matters, urging them to take on management roles to improve the quality of care and patient satisfaction and create a cost-conscious culture (Clark, Spurgeon and Hamilton 2008; Witman et al. 2011; Veronesi, Kirkpatrick and Vallascas 2013).

In this study, a mixed-methods research design is used to explore how the reform of surgical residency training in Netherlands is given shape. The central question was to examine how surgical residents and attending surgeons have implemented and incorporated modern training methodologies, how they respond to external policy reforms and how this affects daily clinical practice. In a more limited and tentative way, the consequences for the evolvement of medical professionalism were examined. As a result, we respond to recent calls to explain the reconfiguration of medical professional work without minimizing the enforcement of organizational control or the strategic resistance of medical doctors to restore professional dominance (Waring and Bishop 2013). This article is based on a collaborative research project (2011–2012) of social scientists working on a national evaluation project on postgraduate medical training reform, and a group of surgeons involved in a regionally organized training program in general surgery. The collaboration between social scientists and practicing surgeons offered the unique opportunity to study the transformation of surgical residency training from the inside out (Dixon-Woods et al. 2011), integrating a more reflective view on (changing) professionalism by exploring the daily experience of medical professionals who have to deal with both traditional and modern expectations and beliefs of postgraduate medical training. To inform the analysis of the empirical findings, the sociological literature on medical education and changing professionalism is discussed in the next paragraphs.

Understanding postgraduate medical training reform

Postgraduate medical training serves 2 main mechanisms underlying professionalism: (1) the succession of medical knowledge and skills, and (2) the transfer and instilling of ethical values and understandings of professional responsibilities (Muzio and Ackroyd 2005). To understand how contemporary reforms have transformed postgraduate medical training, the sociological literature on medical education and theories on changing professionalism was studied, allowing to understand the specific nature and community-based orientation of postgraduate medical training. Insights into changing professionalism, in turn, envision how professional work practices—in our case surgical residency training—have transformed as they increasingly intertwine with organizational structures and purposes (Waring 2014; Blomgren and Waks 2015).

Sociology of medical education

The sociology of medical education, pioneered by Merton, Reader and Kendall (1957) and Becker et al. (1961), is particularly focused on the transition process that medical doctors go through when they enter professional practice. During years of training, they become socialized in the medical professional community by incorporating and constructing a medical identity; taking on the special role and status claimed by the profession (Haas and Shaffir 1982; Hafferty 2000). Through this process they learn to 'think, act, and feel' like a medical doctor (Merton, Reader and Kendall 1957). Hafferty pointed out that construction of this medical identity is a far from straightforward process of replacing one value system by another, but rather is an ongoing and tension-ridden series of encounters during which lay values and attitudes become labeled as 'suspect', 'dysfunctional' and ultimately 'inferior', whereas newly encountered, medical 'ways of seeing and feeling' become internalized as 'desirable', 'functional' and 'superior' (Hafferty 2000: 241–42). Medical identity construction is depicted as a way of both safeguarding professional jurisdictions and autonomy, and organizing work and controlling practitioners (Freidson 1994). Others have focused attention on working in professional practice and the process of embodying expertise. Prentice (2007) and Johnson (2008), for instance, draw on the concept of 'legitimate peripheral participation' (Lave and Wenger 1991) to describe how young physicians come to embody clinical skills, perceptions and judgments by closely working together with their seniors in daily clinical practice. As a result, they gradually become integrated in the local professional community, obtain a more autonomous clinical role and change position from the figurant zone (periphery) to the protected arena (center).

Recently, attention has shifted to the consequence of contemporary reforms in postgraduate medical training, such as the introduction of competency-based curricula and implementation of work hour restrictions, for construction of the medical identity and development of professional autonomy (Bolton, Muzio and Boyd-Quinn 2011; Kellogg 2011; Brooks and Bosk 2012; Wallenburg et al. 2012). In general, these reforms have induced a transformation of the traditional training practice. For example, Brooks and Bosk (2012), have argued that work hour restrictions impinge on the traditional 'rites of passage'; the rituals, tests, ordeals and challenges that must be passed to become an accepted member of the medical community. Among medical professionals, working long hours is traditionally seen as way to enlarge medical knowledge, develop routines, and prove commitment to colleagues. Szymczak et al. (2010) demonstrated that on the one hand work hour restrictions are obeyed, but on the other occasionally are violated when it is felt that these conflict with patient needs or would make them seen as inefficient. Kellogg (2011) has showed that the transformation of postgraduate training programs, which allow medical identity construction the old fashioned way

but are competency-based, structured, transparent and efficient in design, is an increasingly challenging process and can only be achieved by interweaving traditional and modern training methodologies. To better understand how this transformation process unfolds, the literature on changing professionalism is discussed.

Changing professionalism

The literature on changing professionalism encapsulates various theoretical perspectives. Although demarcations are not straightforward or even diffuse, different literature trends can be distinguished. The first, succeeding the classic professionalization literature, has pointed at the de-professionalization of professionals due to increasing outside control. Freidson, for instance, stressed that the professional logic as an ideal type of social organization is increasingly oppressed by the market logic and the managerial logic. He argued that professions 'are seriously weakened in the name of competition and efficiency' (Freidson 2001: 3). Others, drawing on the New Public Management approach, showed how neoliberalism and consumer control have weakened 'classic' professions such as medicine, turning them into occupational professions that primarily face organizational control (Exworthy and Halford 1999).

In the mid-2000s, however, the de-professionalization thesis was increasingly questioned by a growing group of scholars, who argued that de-professionalization actually meant re-professionalization (Duyvendak, Knijn, and Kremer 2006). They stated that professions possess the ability to reframe managerial control (Evetts 2011; Noordegraaf 2011). For example, Levay and Waks (2009) and Waring (2007), demonstrated how the medical profession is able to capture and adapt external regulatory systems to improve quality of care, and consequently can strengthen its position. Other scholars have studied the role of professionals in the process of institutional change (e.g. Scott 2008; Muzio, Brock and Suddaby 2013). They perceived professionals as both the key mechanisms for, and primary targets of institutional change. They also uncovered that professionals have the potential to build, maintain or disrupt institutions (Currie et al. 2012; Muzio, Brock and Suddaby 2013). This concept has drawn attention to the relationship between professionals and institutions, and how professionals are continually engaged in the partial re-enactment of routines and practices, which may ultimately lead to field level dynamism, but may also result in the strengthening of existing institutional arrangements (Lawrence, Suddaby and Leca 2011). Institutional sociologists have particularly highlighted the ability of professions to maintain professional jurisdictions, which show they are able to actively shape institutional arrangements that privilege their own jurisdictional claims and enhance elite status (Currie et al. 2012). Drawing on the work of Lawrence and Suddaby (2006), Currie et al. (2012) uncovered different types of institutional work conducted by professionals: theorizing, defining, educating, policing, constructing normative networks, and finally embedding and routinizing. In another key article in this field, Micelotta and Washington (2013) entitled 'repair work' as another type of institutional work carried out by professionals. This repair work, defined as purposeful and effortful actions to oppose disrupted institutions and re-establish the status quo, is supposed to ensure institutional continuity and stability. Hence, these different literature trends stress the duality between professions and 'external regulators', highlighting the dominance of medical professionals versus non-medical regulators.

The literature on organized professionalism has tried to prevent this division. It has been shown that the organization of expert work is increasingly shaped by a multitude of connections between both the occupational ('professional') and organizational ('managerial') domains (e.g. Noordegraaf 2011; Blomgren and Waks 2015; Postma, Oldenhof and Putters 2015). Literature on organized professionalism has pointed at a shift from the traditional concepts of partnership, collegiality and informality toward a managerial professional business', which is characterized by increasing levels of bureaucracy and formalization (Muzio and Kirkpatrick 2011; Waring and Bishop 2013). This portrays organizations as a key locus and vector of professionalization (Faulconbridge and Muzio 2008; Muzio, Brock and Suddaby 2013). The concept of 'hybrids' is coined to underscore the blurring professional and organizational boundaries, drawing attention to the connections and co-dependencies between professionals and the organizations in which they work (Noordegraaf, van der Steen and van Twist 2014; Waring 2014). Recently, scholars have urged to shift focus from managerial control versus strategic resistance of professionals, to the actual and emergent practices on how professionals interact, mediate and co-create new organizational environments (Waring and Bishop 2013).

In this study, the sociology of medical education is linked to the literature on changing professionalism. Building on this literature and by using a mixed-methods study design, the sweeping transformation of postgraduate medical training in the Netherlands was examined. How do surgeons perceive and deal with contemporary transitions in surgical residency training? How do they give shape to external policy reforms and how do these affect the traditional training practice, and more importantly what potential implications do they have on medical professional evolvement.

METHODS

Setting

In the Netherlands, training in general surgery is organized in 8 training and educational regions, all consisting of 1 university hospital and several affiliated district hospitals. This study was conducted in one of these regions, which is located in the Southwest of the Netherlands. Depending on the training schedule, surgical residents switch between the university hospital and 1 or 2 district hospitals, or vice versa. During their training, surgical residents rotate through various services and departments, gradually taking on increasing levels of autonomy and responsibility. The format and content of the training program is structured in accordance with national requirements set by the Association of Surgeons of the Netherlands and the Central College of Medical Specialists (CCMS).

In 2004, following increasing pressure from the government, the CCMS announced that postgraduate medical training programs had to be reformed to become more patient-centered and transparent in order to meet changing healthcare needs and requirements. According to the resolution, training programs had to be competency-based, with structured clinical rotations and predefined, measurable training objectives (Scheele et al. 2008). In addition, the European Union of Medical Specialists strongly advocated a revision of the duration of postgraduate medical training in order to standardize the widely varying programs between different European countries. As a result, training in general surgery, which often lasted up to 8 years due to postresidency subspecialization fellowships, was shortened to a 6-year program. This significant reduction in training time came on top of the European Working Time Directive, which already limited the average working week to 48 hours and set strict regulations for on-call work.

In response to these developments, a modernized surgical training program based on the Canadian Medical Education Directives for Specialists (CanMEDS) competency framework was introduced in 2009. This framework lists 7 competencies or roles that surgical residents must learn to fulfill, including Medical Expert, Communicator, Collaborator, Manager, Professional, Scholar and Health Advocate (Frank and Danoff 2007). In addition, the format of surgical training was altered. A far more structured training program, comprising clearly defined training objectives, replaced the traditional model of 'learning-by-doing' and 'role modeling'. Various assessment tools were introduced, aiming to standardize assessment and to render performance of competencies both measurable and comparable (Ludmerer and Johns 2005; Lurie, Mooney and Lyness 2009). In addition, subspecialty training was incorporated, with the goal to produce surgeons within a 6-year timeframe who are comparable, in terms of knowledge and skills, to those who completed the former postresidency subspecialization fellowships. Consequently, today's surgical residents are no longer trained as broad-based general surgeons, but as specialists in certain focus areas such as gastrointestinal surgery, surgical oncology, trauma surgery or vascular surgery.

Mixed-methods approach

A mixed-methods study design implies that qualitative and quantitative research methods are used in parallel to inform and strengthen each other, instead of being single components (Creswell and Plano Clark 2007; Goodman et al. 2012). Using both methods iteratively allow the obtained data to be 'mutually informative to construct a negotiated account of what they mean together' (Bryman 2007). In this study, the qualitative data consisted of semi-structured in-depth interviews with surgical residents and attending surgeons, and participant observations during monthly meetings of the regional training committee. This committee, comprising the program directors of the 7 hospitals involved complemented by 1 resident-representative, organizes and monitors surgical residency training on the regional level. In this study, the committee also acted as an expert group with whom the findings were discussed. A questionnaire distributed among all surgical residents and attending surgeons in the training region where the study was conducted provided the quantitative data. Ethical approval for this research was deemed exempt. Below, the qualitative and quantitative components of our mixed-methods study design and their entanglement are described in more detail.

Qualitative methods and analysis

Between July and August 2011, semi-structured interviews with 12 surgical residents and 14 attending surgeons were conducted. In order to obtain a mix of experienced and less routinized, selection was based on seniority. Interview topics were defined on the basis of a previous study on postgraduate medical educational reform by author IW (social scientist), literature research by author CH (PhD-candidate and formal surgical house officer), and practical experience in surgical residency training by the authors PdH and JIJ (both surgeon and program director). Topics included: modernization of the surgical training program, impact of work hour restrictions on surgical residency training, and consequences of subspecialty training for the organization of surgical residency training and healthcare practice. The interviews, conducted by IW, CH and PdH, were tape-recorded and lasted between 30 and 90 minutes and were transcribed verbatim. In addition, IW conducted observations during 7 monthly meetings of the expert group who discussed the national, regional and local training

issues, leading to an overall observation time of 16 hours. Notes taken during the observations were worked out in 'thick descriptions' (Geertz 1973) shortly after.

Analysis of the qualitative data was conducted in two phases. First, IW and CH independently coded the transcribed documents (Atlas. ti 5.0, Cleverbridge, Cologne, Germany) to identify emerging themes and conceptual categories. Deviating codings were discussed between both researchers until consensus was reached. Following this initial stage, CH presented the preliminary results to the expert group to define a framework for the questionnaire. The subsequent discussions among the members of the expert group were observed by IW, and the issues raised were further discussed in the research team (comprising all authors). In general, the expert group meetings, as well as discussions in the research team, created an environment that allowed to obtain different perspectives on the changing landscape of surgical residency training, generating in-depth insight in how the transformation from traditional to modernized training methods is enacted and perceived.

Quantitative methods and analysis

In November 2011, a questionnaire was sent to all surgical residents (n=51) and attending surgeons (n=108), yielding a response rate of 88% (140/159). This questionnaire was also part of study to explore surgeons' attitude toward the CanMEDS competencies and the adopted assessment tools. Detailed information about the questionnaire, as well as demographic data of the respondents can be found in a previously published article by our research team (Hopmans et al. 2013). The questionnaire consisted of the following themes: opinion regarding competency-based training, support and application of assessment tools, subspecialty training and feasibility of training objectives, and impact of work hour restrictions on surgical residency training and healthcare delivery. For each theme, several questions and/or statements were defined, which had to be scored on a 5-point Likert scale: 1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5=strongly agree. Statistical analyses were performed with SPSS version 17.0 (SPSS Inc., Chicago, IL, US). Descriptive statistics were calculated for all items, including means, standard deviations (SD) and frequencies. Agreement with questionnaire items was defined as a response ≥4 on the 5-point Likert scale. To explore differences between the scores of the surgical residents and attending surgeons, the Mann–Whitney U test was applied. The Bonferroni correction was used to adjust p-values for multiple testing, considering $p<0.001$ statistically significant.

Integrating the quantitative and qualitative methods

In this study, both qualitative and quantitative methods were used iteratively to inform eachother, resulting in a mixed-methods analysis. The semi-structured interviews and thick descriptions were used to inform the questionnaire. The statistical analyses of the questionnaire formed the basis for a second analysis of the qualitative data. Subsequently, results of both qualitative and quantitative data were presented at a meeting of the expert group. Observations of the following discussions were used to interpret the outcomes of the questionnaire and connect these to the interview data, leading to insight on how surgical residents and attending surgeons give shape to contemporary transitions in surgical training and how this affects daily clinical practice.

RESULTS

The results are presented in 2 sections. In the first section, we describe how the traditional way of surgical residency training has transformed, and how surgical residents and attending surgeons perceive this change. In the second section, we examine how surgical residents and attending surgeons respond to this transformation in daily clinical practice, namely by conducting repair work to soften unwanted effects and preserve certain values, yet at the same time embrace modern training methodologies to come to a competency-based training program.

Surgical training reform

Historically, surgical residency training has been based on the hierarchal master-apprenticeship model, which is characterized by lengthy periods of on-the-job training under strict supervision of a role model.

In the past, we chose our own role model, someone you wanted to be like. You watched everything he or she – well, usually he – did. You joined him at the outpatient clinic, watched his activities on the ward, and accompanied him in the operating theater. You wanted to see and learn as much as you could, meaning you started at 7:00 AM and stayed until 20:00-21:00 PM. Nowadays, residents are less present. That's why reform of surgical training is necessary. You need formal training objectives and assessment of skills to know how someone is doing (attending surgeon [AS]11).

As quoted above, the attending surgeon points at the crucial importance of role modeling, which can be defined as closely interacting with and 'shadowing' of a respected surgeon

whose behavior and skills are admired. Due to this close cooperation, as well as working long hours, performance and progress was easily to monitor. Yet, this attending surgeon argues that the sharp reduction in the number of hours surgical residents are allowed to spend in the hospital encroach on the traditional master-apprenticeship model, underlining and recognizing that training reform is inevitable.

An important part of surgical training reform involved the implementation of a modern training methodology. During the interviews, the majority of the surgical residents and attending surgeons expressed their support for the competency-based training and assessment program. However, quantitative data revealed a different opinion (Table 1). Although most surgical residents (77%, mean=3.91) and attending surgeons (83%, mean=3.90) indicated to be aware of the implementation of a modern training methodology, a minority of both groups considered it an appropriate way of training (surgical residents 44%, mean=3.35 versus attending surgeons 29%, mean=3.21). Moreover, both surgical residents and attending surgeons still highly appreciated the master-apprenticeship model (surgical residents 86%, mean=4.26 versus attending surgeons 78%, mean=3.89). Not surprisingly, the competency-based training program was not considered an improvement as compared to the traditional master-apprenticeship model (surgical residents 14%, mean=2.93 versus attending surgeons 21%, mean=2.93).

The incorporation of subspecialty training is another major change. After 4 years of training in general surgery, surgical residents must choose a subspecialty on which they concentrate during the last 2 years of their training. Consequently, surgical residents are no longer trained as general surgeons, being able to diagnose, treat, and manage patients with a broad spectrum of surgical conditions, but as specialists who are qualified in a certain focus area. As shown in Table 2, most of the respondents recognized the need for subspecialty training. In addition, most surgical residents (75%, mean=3.77) and attending surgeons (73%, mean=3.97) believed that subspecialty training will improve the quality of patient care. During the interviews, it was argued that due to innovations in medical technology and the constantly expanding field of medicine, it is no longer possible for a single individual to have up-to-date knowledge of all aspects of a surgical discipline and stay competent to perform a wide range of surgical procedures.

You don't want a vascular surgeon to operate on your mother's broken hip. He's less competent than a trauma surgeon who performs multiple hip operations a week. That's just the way it is (surgical resident [SR]1).

Table 1. Summary of questionnaire data on respondents' agreement with statements regarding the implementation of a modernized surgical training program.

	Surgical Residents				Attending Surgeons				
Statements	**N**	**agreement [a] (%)**	**mean (/5)**	**SD**	**N**	**agreement [a] (%)**	**mean (/5)**	**SD**	**p-value**
I am aware of the implementation of a modernized training program	43	77	3.91	0.75	97	83	3.90	0.76	0.899
The modernized program is an appropriate training model	43	44	3.35	0.72	97	29	3.21	0.64	0.217
The master-apprenticeship model forms an essential part of surgical residency training	43	86	4.26	0.69	97	78	3.89	0.87	0.021
The modernized training program is an improvement compared to the traditional master-apprenticeship model	43	14	2.93	0.67	97	22	2.93	0.84	0.907

[a] Agreement is defined as a score ≥4 (indicating a positive attitude) on the 5-point Likert-scale

However, both qualitative and quantitative data also reflected concern about the shift to subspecialty training, especially regarding the consequences that it might have on the organization of surgical healthcare and surgical residency training. As demonstrated in Table 2, 47% (mean=3.33) of the surgical residents and 63% (mean=3.59) of the attending surgeons perceived that general surgeons will remain needed in the future. Almost a similar percentage (surgical residents 48%, mean=3.26 versus attending surgeons 51%, mean=3.34) indicated that general surgery should be added as a subspecialty in the training program. Furthermore, both groups were uncertain whether it would be feasible to train qualified surgical specialists within 6-year timeframe who are comparable to surgeons who completed the former postresidency subspecialization fellowships (total training period 8 years), with agreement percentages ranging from 37 to 56% for the different subspecialties offered. In addition, only 14% (mean=2.26) of the surgical residents and 11% (mean=2.16) of the attending surgeons agreed with the statement that postresidency subspecialization fellowships will be superfluous in the future, indicating a general belief that surgical residents are not ready to act as independent, fully competent surgeons in their field of expertise after completion of the 6-year training program (Table 2).

During the interviews, another concern, which could be related to the shift to tailor-made subspecialty training schemes, was noticed. As not all surgical procedures are carried out in every hospital anymore, especially not the more complicated procedures such as lung surgery or particular kinds of cancer operations, some surgical residents now have to compose an individual training program, which requires them to work in different hospitals in order to meet the training objectives for their chosen subspecialty. Both surgical residents and attending surgeons worried about the consequences of 'moving and shopping around'. The lack of opportunity to build a personal relationship was of particular concern. During rotations through various surgical services and departments, surgical residents gradually become part of the team and are allowed to take on an increasing level of responsibility and autonomy. Frequently moving in and out of a particular hospital as a consequence of tailor-made subspecialty training schemes, or due to an increase in irregular hours of shift-work resulting from the implementation of the European Working Time Directive, induces social distance between surgical residents and attending surgeons:

Previously, residents saw everything. Now, the day is divided into several shifts. Therefore, there is lack of continuity. Attending surgeons work more hours than surgical residents do. Especially when you occupy a lower level in hierarchy, you lack information. When you have a more prominent position, you'll be informed; otherwise you have to sort things out for yourself. For educational purposes, it would be better if they were the ones solving the problems instead of us (AS 3).

Table 2. Summary of questionnaire data on respondents' agreement with statements regarding the feasibility of training objectives and subspecialty training.

	Surgical Residents				Attending Surgeons			
Statements	**N**	**agreement [a] (%)**	**mean (/5)**	**SD**	**N**	**agreement [a] (%)**	**mean (/5)**	**SD**	**P-value**
Subspecialty training will improve the quality of patient care	43	75	3.77	1.04	95	73	3.97	0.88	0.41
General surgeons remain needed in the future	43	47	3.33	1.04	95	63	3.59	1.14	0.132
General surgery should be added as a subspeciality in the modernized training program	43	48	3.26	1.04	94	51	3.34	1.19	0.646
Vascular surgeons can be trained in 6 years	43	44	3.12	1.14	94	47	3.19	1.13	0.718
Trauma surgeons can be trained in 6 years	43	54	3.42	1.12	94	56	3.38	1.06	0.868
Gastrointestinal surgeon can be trained in 6 years	43	37	3.02	1.11	94	49	3.20	1.08	0.351
Surgical oncologists can be trained in 6 years	43	42	3.09	1.06	94	51	3.30	1.09	0.286
Postresidency subspecialization fellowships will be superfluous in the future	43	14	2.26	0.98	94	11	2.16	0.97	0.560

[a] Agreement is defined as a score ≥4 (indicating a positive attitude) on the 5-point Likert-scale

During their training, surgical residents are part of a vested partnership of surgeons for only a short period of time. A surgeon typically pointed them out as 'passers by'. Hence, surgical residents lack the informal connections which attending surgeons do have. Attending surgeons, for instance, are informed and being kept up to date about patients admitted to the wards after a day off or a holiday. As surgical residents lack these informal connections, it is difficult for them to stay closely involved and thus participate in daily clinical practice. Moreover, moving between different hospitals sharply contrasts with the tradition of gradual integration and embedded learning.

I don't think it's possible to learn a specific trick [surgical procedure] within a year or so. Look, you enter someone else's shop, and it takes months before a boss [attending surgeon] really trusts you. Before they allow you to perform a surgical procedure independently, a year has passed (SR 5).

Imagine, a resident wants to learn a specific form of lung surgery in my hospital for a period of 3 months. Then the others will say, here he comes, pinching our operations. That's not going to work (AS 2).

These interview excerpts underscore and reinforce earlier accounts on surgical residency training that entrance to the core, which is the operating theater, needs to be earned by demonstrating competence, dedication and seniority. In addition, surgical residents must put effort in becoming a 'team member' and build up trust and respect in the group of surgical residents. The tailor-made subspecialty programs with a focus on the individual instead of group learning impede on the process of embedded learning and socialization.

Implementation of the European Working Time Directive has not only limited the number of hours surgical residents are allowed to work per week, but also sets requirements for weekly rest periods and stipulates regulations for night and shift-work. During the interviews and monthly meetings of the expert group, these work hour regulations were a frequent, and sometimes heated, topic of debate. It was generally believed that work hour restrictions impinge on surgical residency training. This belief is supported by questionnaire data, which show that a minority of the surgical residents (47%, mean=3.09) indicated that a 48-hour working week provides enough time for adequate surgical training. In the group of attending surgeons, this percentage decreased to 26% (mean=2.76), indicating a perception that a decrease in working hours comes at the expense of the quality of training. Other statements exploring the impact of work hour regulations demonstrate that attending surgeons are significantly more concerned about the consequences of the 48-hour working week with regard to the continuity of surgical residency training (p<0.001), health care delivery (p<0.001), and the number of clinical patient encounters by surgical residents (p<0.001). These results can be seen in Table 3.

Table 3. Summary of questionnaire data on respondents' agreement with statements regarding the impact of work hour restrictions.

	Surgical Residents				Attending Surgeons				
Statements	**N**	**agreement [a] (%)**	**mean (/5)**	**SD**	**N**	**agreement [a] (%)**	**mean (/5)**	**SD**	**p-value**
A 48-hour working week provides sufficient time for adequate surgical training	43	47	3.09	1.02	95	26	2.76	0.94	0.051
Work hour restrictions negatively affect the continuity of surgical residency training	43	49	3.35	1.19	95	84	3.99	0.77	<0.001
Restrictive work hour regulations limit the number of surgical procedures performed by residents	43	68	3.65	0.92	94	78	3.85	0.88	0.199
Work hour restrictions decrease the number of clinical patient encounters by residents	43	49	3.30	1.08	95	83	3.95	0.80	<0.001
Restrictive work hour regulations negatively affect the continuity of health care delivery	43	56	3.49	1.08	94	92	4.16	0.59	<0.001

[a] Agreement is defined as a score ≥4 (indicating a positive attitude) on the 5-point Likert-scale

Concerns on a decreased clinical exposure were not only due to the substantial reduction in working hours, but also a result of an increasing emphasis on efficiency in hospital organizations, particularly in the use of the operation theater. The operating theater is a costly part of a hospital organization because of the use of high-tech medical equipment and presence of highly educated staff. To save time and money, surgical procedures are accurately scheduled and time is managed strictly. In some hospitals, time management policies have stimulated discussion on the enrollment of surgical residents in operations as they need strict guidance and instruction and hence work in a far less efficient way than routinized attending surgeons. As one of the surgeons explained:

If you do not complete the surgical procedures within the assigned timeframe, you do not manage to keep up with the production. Sometimes this comes at the cost of surgical residents' operating time. For example, at the end of a day when there is pressure of time, some operating room managers allow you to perform an operation, but only when you do it yourself (AS 12).

Although not all attending surgeons shared this experience, time for training cannot be taken for granted and increasingly must be negotiated. An increased focus on the strategic policy of hospitals increasingly depend on organizational properties that serve both professional and organizational aims: providing good care and working in an efficient manner, as well as training a new generation of surgeons and holding on to operation schedules. Hence, external regulations and organizational preferences increasingly must be taken into account.

In the next section, we demonstrate how surgical residents and attending surgeons have responded to the sweeping transformation surgical residency training by conducting repair work. This repair work encapsulates both implementing modern training methodologies and adapting to external policy reforms while simultaneously seeking convenient and feasible solutions to preserve certain values and related practices

Repairing reforms

From the data analysis, 3 repertoires of repair work emerged: organizing the reform, bending rules, and negotiating instruments.

Organizing the reform

An important part of the surgical training reform entails a shift from the master-apprenticeship model to a structured, competency-based training and assessment program. The excerpt below illustrates how the attending surgeons 'organized the reform', revealing their support to make it a success.

The surgical program directors discuss the training schemes that are displayed in colored charts in front of them. Some residents are still looking for a training post in the subspecialty of their choice. One of the program directors complains that he is losing residents to a neighboring hospital, while his group invested so much in training them. The chair program director answers that they all face this problem, but that the training schemes of the residents are leading now. The case of Tom [fictional name] is discussed. He has signed up for subspecialty training in vascular surgery. The chair program director glances at one of the others: can you have him? I don't know him personally, he answers. I want to talk to him first. Tom is paged and instructed to show up at the end of the meeting (field note made during meeting of the expert group).

This excerpt demonstrates how the program directors strive to compose training schemes according to the desires of individual surgical residents. This process is not without its setbacks and difficulties as some hospitals 'lose' their experienced and hence valuable surgical residents to neighboring hospitals. Fabricating tailor-made schemes and negotiating the allocation of surgical residents is extremely time consuming. As shown, personal relationships play a pivotal role. Program directors would like to know who they have got to work with, illustrating that in addition to the aim to facilitate training, personal relationships and trust continue to be of paramount performance.

As demonstrated in Table 3, surgical residents were particularly concerned about the impact of the work hour restrictions on their operative caseload, with 68% (mean=3.65) indicating that work hour restrictions have limited the number of surgical procedures performed. In daily clinical practice, however, surgical residents ensure that the operative caseload they are exposed to is preserved by dividing operations among the group of surgical residents taking into account the level of competence of each individual.

You can always discuss your scheduled operations with one of the senior residents. For example, if someone has done more than enough gallbladder operations, these can be passed on to another person (SR 7).

Surgical residents thus conduct repair work by scheduling specific surgical procedures for individuals. This allows to practice particular surgical skills and organize learning opportunities according to individual needs.

Bending rules

A second repertoire of repair work is 'bending rules'. The qualitative data revealed that both surgical residents and attending surgeons sometimes ignore regulations to create training opportunities or preserve 'old school' training methods, like 'sticking to the ward' instead of going home after a shift. As one of the surgical residents pointed out:

Hanging around. Sticking to the ward. If you want to operate, you need to be the first. It is always the same group of residents, who perform most operations. Those who stand out, get rewarded. They are phoned more often than the residents who don't hang around. This is an important but underestimated part of surgical training (SR 4).

Demonstrating eagerness to learn and, in doing so, indirectly achieve bonding with the group of attending surgeons, is a way of creating training opportunities (Wallenburg, Pols and de Bont 2015). Another surgical resident explained sometimes to not register his/her name in order to perform a specific surgical procedure, making presence invisible to external regulators when working overtime. In the Netherlands, hospitals risk high fines for violating work hour restrictions. Not registering, as way of working around the system, may be a convenient solution to maximize training opportunities or meet specific training objectives:

I want to become a good surgeon. This means that I must be able to remove a spleen. These patients do not show during office hours. They enter in the middle of the night (SR 5).

These interview excerpts parallel Szymczak and colleagues' argument that surgical residents do not simply disregard work hour restrictions, but rather bend the rules by seeking suitable solutions to local or individual needs (Szymczak et al. 2010). This repair work is aimed to soften or even solve the perceived detrimental effect of the work hour restrictions on the quality of postgraduate medical training.

Negotiating instruments

A third repertoire of repair work is the negotiation of instruments. With the implementation of competency-based training, assessment tools have been adopted to evaluate surgical residents' development and progression in attaining the various roles of the CanMEDS competency framework. Briefly, these assessment tools are standardized checklists aimed at making competence progression objective, measurable and comparable (Epstein 2007). For example, the instrument Objective Structured Assessment of Technical Skills (OSATS) is used to assess different dimensions that are important when performing a surgical procedure, including respect for tissue, time and motion, knowledge and handling of instruments, use of assistants, flow of operation and knowledge of the specific procedure.

OSATS are great. They provide structure and tell at what level you're performing. They also force attending surgeons to interact with you and provide feedback on how you're doing (SR 9).

This interview excerpt shows that the assessment tools allow surgical residents to gain grip on their learning process and can help to obtain feedback and personal attention. Similar results emerged from the quantitative data. Most of the respondents (surgical residents 79%, mean=3.81 versus attending surgeons 73%, mean 3.94) agreed with that assessment tools are used to structure and coordinate feedback (Table 4). Surprisingly, the qualitative data revealed that attending surgeons appreciate the assessment tools for different reasons by using them to detect underperformance or stimulate adequately functioning surgical residents to excel, as pointed out below.

The benefit of competency-based training is that the assessment tools act as a safety net and provide tangible evidence when a resident is functioning substandardly compared to his/her peers (AS 1)

When evaluating residents [using assessment tools], I always make sure to point out something that can be improved. Scoring everything at the top level makes no sense (AS 12)

Interestingly, during the interviews the use of assessment tools was particularly related to surgical residents in their early days of training and to those who did not seem to meet training objectives and/or professional norms. Conversely, adequately performing experienced surgical residents seemed to have passed beyond the assessment tools. During our research period, it was noted that the assessment tools became less significant and lessons were transferred more implicitly during personal interactions. One of the attending surgeons states:

You have to make clear what can be improved, and that's not something you trust to paper but tell someone in person. Such feedback is far more important than circling grades [on the assessment tools]. I don't care about these grades. In the end, you judge residents by comparing them to yourself. Setting norms is pointless; it's all about trust (field note made during meeting of the expert group).

This remark is supported by questionnaire results, demonstrating that most surgical residents (72%, mean=3.98) and attending surgeons (71%, mean=3.83) perceived that assessment tools cannot replace attending surgeons' opinion of a surgical resident. Thus, professional intuition is considered superior to circled grades on 'objective' assessment tools. In addition,

Table 4. Summary of questionnaire data on respondents' agreement with statements regarding the support and application of assessment tools.

	Surgical Residents				Attending Surgeons				
Statements	**N**	**agreement [a] (%)**	**mean (/5)**	**SD**	**N**	**agreement [a] (%)**	**mean (/5)**	**SD**	**p-value**
Assessment tools are used to structure and coordinate feedback	43	79	3.81	0.76	95	73	3.94	0.60	0.522
Assessment tools cause an administrative burden	43	81	4.02	0.83	96	86	4.20	0.72	0.285
Assessment tools cannot replace attending surgeons' opinion of a resident	43	72	3.98	0.91	96	71	3.83	0.91	0.397

[a] Agreement is defined as a score ≥4 (indicating a positive attitude) on the 5-point Likert-scale

the assessment tools were blamed for causing an administrative burden by 81% (mean=4.02) of the surgical residents and 86% (mean=4.20) of the attending surgeons. These results can be seen in Table 4. Hence, it can be concluded that the introduction of assessment tools has not completely replaced the subjective evaluations from the master-apprenticeship training model, uncovering tension between a formalized and objective way of competence testing on the one hand, and reliance on personal interactions and professional judgment on the other. However, in this study, assessment tools were repaired by embedding them in professional interactions during daily clinical practice, allowing them to become part of the ongoing socialization process of becoming a surgeon.

DISCUSSION

In this study, we focused on the critical case of postgraduate medical training to examine how medical professionals give shape to modern training methodologies and external policy reforms, and how this impacts on training in daily clinical practice. Using a mixed-methods analysis, we revealed that both surgical residents and attending surgeons are rather ambivalent toward the reform. On the one hand, the increased structure and formalization of surgical residency training is perceived as a necessary answer to external policy reforms (i.e. restrictive work hour restrictions, shortening the duration of the training program) due to which less time is spent in the hospital and personal contact between surgical resident and attending surgeon has decreased. Yet, on the other hand the transformation of surgical residency training is being criticized as it impedes on highly valued traditions, such as the gradual integration and socialization in a medical community. In addition, it impinges on established educational principles like role modeling and learning by doing.

We have shown how surgical residents and attending surgeons give meaning to modern training methodologies and external policy reforms by simultaneously implementing and 'repairing' them in daily clinical practice. This repair work aims to soften the unwanted effects and re-embed surgical residency training in the social-material context of professional practice. Repair work, in this study, involved organizing for the reform, bending rules and negotiating instruments. For instance, program directors put a lot of effort in establishing individual training schemes and sought to maintain personal relationships with the surgical

residents. Surgical residents, in turn, collaborate in organizing their participation in the operating theater. In doing so, both the surgical residents and attending surgeons connect modern training methodologies with established practices of embedded training and socialization.

Repair work, we argue, goes beyond restoring disrupted institutional arrangements to restore the status quo in the sense of re-establishing the old (Micelotta and Washington 2013), mirroring the scholarly debate of professions as controlled by managerial processes or strategically operating actors resisting change. Instead, it is an interactive process; implementing modern training methodologies and adapting to external policy reforms while simultaneously seeking convenient and feasible solutions to preserve certain values and related practices. Hence, repair work encapsulates a more diverse variety of activities of which outcomes are yet unsure and emergent. This study contributes to the developing institutionalist perspective on professionalism, and more in particular to the institutional work literature (e.g. Currie et al. 2012; Lawrence et al. 2013; Micelotta and Washington 2013). Repair work, as a specific type of institutional work, is heterogeneous in nature. It is an ongoing and recursive process of goal-setting, learning and revision that is based on learning that impacts on vested practices, prompting institutional reconfiguration. As a result, routines emerge that reflect shifting institutional relationships; surgical residency training no longer is an isolated practice merely using the operating theater. Program directors continuously must negotiate the use of other hospital facilities, underlining how expert work is increasingly situated in the bureaucratic hospital practice.

In addition, and interesting in the light of recent literature on the role of standardization and materiality in institutional work (e.g. Slager, Gond and Moon 2012; Monteiro and Nicolini 2015) that is focused on how procedures and objects are used to perform institutional work, we demonstrated how material entities (e.g. assessment tools, training schemes) mediate and pursue institutional change and transform vested professional practices and daily routines. Establishing tailor-made training schemes, for instance, puts the desires and learning goals of individual surgical residents at the heart of the training process, hence encroaching on the traditional hierarchical master-apprentice relationship. Material entities thus play an active role in constituting and perpetuating postgraduate medical training programs. Future research should further explore the role of materiality in the evolvement of postgraduate medical training, potentially generating additional insights in the actual day-to-day actions and processes that underlie medical professional evolvement more in general.

CONCLUSIONS

Medical education has been said to be stubborn to change. In a special issue on medical education in the Journal of Health and Social Behavior in 1988, Bloom stated: 'education is essentially unchanging, even though brave ideological statements guide curriculum reforms that do little but mask the underlying reality' (Bloom 1988: 295). This study, however, has demonstrated the ability of medical professionals – in our case surgical residents and attending surgeons involved in a regionally organized training program in general surgery- to renew and adapt daily routines to continuously changing societal and organizational circumstances, sketching a more dynamic and reflexive picture of medical professional evolvement. We have shown how change processes unfold as medical doctors give meaning to modern training methodologies and external policy reforms by simultaneously implementing and 'repairing' them in daily clinical practice, and how this reconfigures medical professionalism.

REFERENCES

Becker, H. S. et al. (1961) Boys in White: Student Culture in Medical School (fourth printing 1992). New Brunswick and London: Transaction Publishers.

Blomgren, M., and Waks, C. (2015) 'Coping with Contradictions: Hybrid Professionals Managing Institutional Complexity', Journal of Professions and Organization, 2: 78–102.

Bloom, S. W. (1988) 'Structure and Ideology in Medical Education: An Analysis of Resistance to Change', Journal of Health and Social Behavior, 29: 294–306.

Bolton, S. C., Muzio, D., and Boyd-Quinn, C. (2011) 'Making Sense of Modern Medical Careers: The Case of the UK's National Health Service', Sociology, 45: 682–99.

Bosk, C. L. (1979) Forgive and Remember: Managing Medical Failure. Chicago: The University Press of Chicago Press.

Brooks, V. J., and Bosk, C. L. (2012) 'Remaking Surgical Socialization: Work Hour Restrictions, Rites of Passage and Occupational Identity', Social Science & Medicine, 75: 7525–32.

Bryman, A. (2007) 'Barriers to Integrating Quantitative and Qualitative Research', Journal of Mixed Methods Research, 1: 8–22.

Clark, J., Spurgeon, J., and Hamilton, P. (2008) 'Medical Professionalism: Leadership Competency–An Essential Ingredient', The International Journal of Clinical Leadership, 16: 3–9.

Creswell, J. W., and Plano Clark, V. L. (2007) Designing and Conducting mixed Methods Research.Thousand Oaks: Sage.

Currie, G. et al. (2012) 'Institutional Work to Maintain Professional Power: Recreating the Model of Medical Professionalism', Organization Studies, 33: 937–62.
—— Finn, R., and Martin, G. (2009) 'Professional Competition and Modernizing the Clinical Workforce in the NHS', Work, Employment and Society, 23: 267–84.

Dixon-Woods, M. et al. (2011) 'Explaining Michigan: Developing an Ex Post Theory of a Quality Improvement Program', The Milbank Quarterly, 89: 167–205.

Duyvendak, J., Knijn, W. T., and Kremer, M. (2006) 'Policy, People and the New Professional: An Introduction', in Knijn, T., Kremer, M., and Duyvendak, J. W. (eds) Policy, People and the New Professional: De-Professionalisation and Re-Professionalisation in Care and Welfare, pp. 19–33. Amsterdam: Amsterdam University Press.

Epstein, R. M. (2007) 'Assesment in Medical Education', The New England Journal of Medicine, 356: 387–96.

Escobar, M. A., and McCullough, L. B. (2006) 'Responsibly Managing Ethical Challenges of Residency Training: A Guide for Surgery Residents, Educators and Residency Program Leaders', Journal of American College of Surgeons, 202: 531–35.

Evetts, J. (2011) 'A New Professionalism? Challenges and Opportunities', Current Sociology, 59: 406–22.

Exworthy, M., and Halford, S., eds. (1999) Professionals and the New Managerialism in the Public Sector, Buckingham, United Kingdom: Open University Press.

Faulconbridge, J., and Muzio, D. (2008) 'Organizational Professionalism in Globalizing Law Firms', Work, Employment and Society, 22: 7–25.

Frank, J. R., and Danoff, D. (2007) 'The CanMEDS Initiative: Implementing an Outcomes-Based Framework of Physician Competencies', Medical Teacher, 29: 642–47.

Freidson, E. (1994) Professionalism Reborn: Theory, Propercy and Policy. Chicago: Chicago University Press.
—— (2001) Professionalism: The Third Logic. Cambridge, United Kingdom: Polity.
Geertz, C. (1973) The Interpretation of Cultures. New York: Basic Books.

Goodman, A. et al. (2012) 'Healthy Travel and the Socio- Economic Structure of Care Commuting in Cambridge, UK: A Mixed-Methods Analysis', Social Science & Medicine, 74: 1929–38.

Haas, J., and Shaffir, W. (1982) 'The Hidden Curriculum of Professionalization', Work and Occupations, 9: 131–54.

Hafferty, F. W. (2000) 'Reconfiguring the Sociology of Medical Education: Emerging Topics and Pressing Issues', in Bird, C. E., Conrad, P., and Fremont, A. M. (eds) Handbook of Medical Sociology, 5th edn. pp. 238–57. New Jersey: Prentice Hall.

Hopmans, C.J., et al. (2013) 'Surgeons' attitude toward a competency- based training and assessment program: results of a multicenter survey', Journal of Surgical Education, 70(5): 647–54.

Johnson, E. S. (2008) 'Out of My Viewfinder, Yet in the Picture: Seeing the Hospital in Medical Simulations', Science, Technology & Human Values, 33: 53–76.

Kellogg, K. C. (2011) 'Operating Room: Relational Spaces and Microinstitutional Change in Surgery', American Journal of Sociology, 115: 657–711.

Kipping, M., and Kirkpatrick, I. (2013) 'Alternative Pathways of Change in Professional Service Firms: The Case of Management Consulting', Journal of Management Studies, 50: 777–807.

Lave, J., and Wenger, E. (1991) Situated Learning Legitimate Peripheral Participation. Cambridge: Cambridge University Press.

Lawrence, T. B., and Suddaby, R. (2006) 'Institutions and Institutional Work', in Clegg, S. R. et al. (eds) Handbook of Organization Studies, 2nd edn. pp. 215–54. London: Sage.
—— Leca, B. (2011) 'Institutional Work: Refocusing Studies of Organization', Journal of Management Inquiry, 20: 52–8.
——Suddaby, R., Leca, B., and Zilber, T. B. (2013) 'Institutional Work: Current Reserach, New Directions and Overlooked Issues', Organization Studies, 34: 1023–33.

Levay, C., and Waks, C. (2009) 'Professions and the Pursuit of Transparency in Healthcare: Two Cases of Soft Autonomy', Organization Studies, 30: 509–27.

Ludmerer, K. M., and Johns, M. M. E. (2005) 'Reforming Graduate Medical Education', Journal of American Medical Association, 294: 1083–87.

Lurie, S. J., Mooney, C. J., and Lyness, J. M. (2009) 'Measurement of the General Competencies of the Accreditation Council for Graduate Medical Education: A Systematic Review', Academic Medicine, 84: 301–09.

McGivern, G. et al. (2015) 'Hybrid Manager-Professionals' Identity Work: The Maintenance and Hybridization of Medical Professionalism in Managerial Contexts', Public Administration, 93: 412–32.

Merton, R. K., Reader, G. G., and Kendall, P. L. (1957) The Student-Physician:Introductory Studies in the Sociology of Medical Education. Cambridge, MA: Harvard University Press.

Micelotta, E. R., and Washington, M. (2013) 'Institutions and Maintenance: The Repair Work of Italian Professions', Organization Studies, 34: 1137–70.

Monteiro, P., and Nicolini, D. (2015) 'Recovering Materiality in Institutional Work: Prizes as an Assemblage of Human and Material Entities', Journal of Management Inquiry, 24: 61–81.

Muzio, D., and Ackroyd, S. (2005) 'On the Consequences of Defensive Professionalism: Recent Changes in the Legal Labour Process', Journal of Law and Society, 32: 615–42.
——and Kirkpatrick, I. (2011) 'Introduction: Professions and Organizations: a Conceptual Framework', Current Sociology, 59: 389–405.
——Brock, D. M., and Suddaby, R. (2013) 'Professions and Institutional Change: Towards an Institutionalist Sociology of Professions', Journal of Management Studies, 50: 699–721.

Noordegraaf, M. (2007) 'From "Pure" to "Hybrid" Professionalism. Present-Day Professionalism in Ambiguous Public Domains', Administration & Society, 39: 761–85.
——(2011) 'Risky Business: How Professionals and Professional Fields (Must) Deal with Organizational Issues', Organization Studies, 32: 1349–71.
——,van der Steen, M., and van Twist, M. (2014) 'Fragmented or Connective Professionalism? Strategies for professionalizing the work of strategists and other (Organizational) Professionals', Public Administration, 92: 21–38.

Postma, J., Oldenhof, L., and Putters, K. (2015) 'Organized Professionalism in Healthcare: Articulation Work by Neighbourhood Nurses', Journal of Professions and Organization, 2: 61–77.

Prentice, R. (2007) 'Drilling Surgeons: The Social Lessons of Embodied Surgical Learning', Science, Technology & Human Values, 32: 534–53.

Ringsted, C. et al. (2006) 'Are some of the challenging aspects of the canmeds roles valid outside Canada?', Medical Education, 40: 807–15.

Scheele, F. et al. (2008) 'Introducing Competency-Based Postgraduate Medical Education in the Netherlands', Medical Teacher, 30: 248–53.

Scott, W. R. (2008) 'Lords of the Dance: Professionals as Institutional Agents', Organization Studies, 29: 219–38.

Slager, R., Gond, J. P., and Moon, J. (2012) 'Standardization as Institutional Work: The Regulatory Power of a Responsible Investment Standard', Organization Studies, 33: 763–90.

Szymczak, J. E. et al. (2010) 'To Leave or Lie? Are Concerns About a Shift-Work Mentality and Eroding Professionalism as a Result of Duty-Hour Rules Justified?', The Milbank Quaterly, 88: 350–81.

Veronesi, G., Kirkpatrick, I., and Vallascas, F. (2013) 'Clinicians on Board: What Difference Does it Make?', Social Science & Medicine, 77: 147–55.

Wallenburg, I., Pols, J., and de Bont, A. (2015) '"You need to bond with the ones you train": Mixing epistemic cultures in resindency training', Evidence & Policy, 11(3): 399–414.
——et al. (2013) 'Learning to Doctor: Tinkering with Visibility in Residency Training', Sociology of health & Illness, 35(4): 544–59.
——et al. (2012) 'Negotiated authority: A comparative study of reform in medical training regimes', Journal of Health Politics, Policy and Law, 37(3): 339–67.
——et al. (2010) 'Between trust and accountability: different perspectives on the modernization of medical specialist training in the Netherlands', Academic Medicine, 85(6): 1082–90.

Waring, J. (2007) 'Adaptive Regulation or Governmentality: Patient Safety and the Changing Regulation of Medicine', Sociology of Health & Illness, 29: 163–79; 63.
——(2014) 'Restratification, Hybridity and Professional Elites: Questions of Power, Identity and Relational Contingency at the Points of Professional-Organisational Intersection', Sociology Compass, 8: 688–704.
——and Bishop, S. (2013) 'McDonaldization or Commercial Re-Stratification: Corporatization and the Multimodel of English Doctors', Social Science & Medicine, 82: 147–55.
—— and Currie, G. (2009) 'Managing Expert Knowledge: Organizational Challenges and Managerial Futures for the UK Medical Profession', Organization Studies, 30: 755–78.

Witman, Y. et al. (2011) 'Doctors in the Lead: Balancing Between Two Worlds', Organization, 18: 477–95.

Chapter 4

Assessment of surgical residents' operative skills in the operating theater using a modified Objective Structured Assessment of Technical Skills (OSATS): A prospective multicenter study

Cornelis J. Hopmans
Pieter T. den Hoed
Lijckle van der Laan
Erwin van der Harst
Maarten van der Elst
Guido H.H. Mannaerts
Imro Dawson
Reinier Timman
Bas P.L. Wijnhoven
Jan N.M. IJzermans

Adapted from: Surgery. 2014 Nov;156(5):1078-88

ABSTRACT

Background: With the implementation of competency-based curricula, Objective Structured Assessment of Technical Skills (OSATS) increasingly is being used for the assessment of operative skills. Although evidence for the usefulness of this instrument has been demonstrated in experimental study designs, data supporting OSATS application in the operating room are limited. This study evaluates the validity and reliability of the OSATS instrument to assess surgical residents' operative skills in the operating theater.

Methods: Twenty-four surgical residents were recruited from 7 hospitals within a surgical training region and classified equally into 3 groups according to postgraduate training year (PGY). Each participant had to perform 5 different types of operations. Surgical performance was measured using a modified OSATS instrument consisting of 3 scales: Global Rating Scale, Overall Performance Scale and Alphabetic Summary Scale. Validity and reliability metrics included: construct validity (Kruskal-Wallis test) and internal consistency reliability (Cronbach's α coefficient). Spearman's correlation coefficients were calculated to determine correlations between the different scales.

Results: Eighteen surgical residents [PGY1-2 (n=7); PGY3-4 (n=8); PGY5-6 (n=3)] performed 249 operations. Comparisons of the performance scores revealed that evidence for construct validity depended on the difficulty level of the selected procedures. For individual operations, internal consistency reliability of the Global Rating Scale ranged from 0.93-0.95. Scores on the different scales correlated strongly ($r=0.62-0.76$, $p<0.001$).

Conclusion: Assessment of operative skills in the operating theater using this modified OSATS instrument has the potential to establish learning curves, allowing adequate monitoring of surgical residents' progress in achieving operative competence. Use of the Overall Performance Scale should be reconsidered. The Alphabetic Summary Scale seems to be of additional value for clinical practice.

INTRODUCTION

Traditionally, surgical residency training has been based on the master-apprenticeship model. In this training model, characterized by strict obedience to and copying of a role model, operative skills were almost exclusively learned in the operating room by working long hours and participating in as many operations as possible. However, this way of training has been criticized for the subjective assessment of surgical residents' skills, often based on surrogate markers of operative competence, such as accumulated case numbers and time taken for a procedure, or biased by personal preferences of supervising surgeons.[1,2] In addition, the implementation of work hour restrictions has substantially reduced the available training time, leaving less time to be spent in the operating theater.[3] As a result, training opportunities are limited and operative competence based on excessive exposure to a large variety of operations can no longer be assumed.[4] These developments, in combination with an increased emphasis on debates concerning patient safety and a greater demand for accountability and transparency in medicine, from both a professional and public point of view, have initiated a shift toward a more objective and formalized approach of postgraduate medical training.[5] In response, many Western countries have recently restructured the format of their surgical training programs into competency-based curricula.[6-9] In these curricula there is a strong need for educational tools that are capable of providing objective performance assessments to evaluate predefined competency goals.[10]

Although surgical skills represent only a part of the qualities needed to become a well-trained surgeon, development of operative competence plays a pivotal role in any surgical training program.[11] Within the past decades, various assessment methods have been developed for the evaluation of operative skills, including observational instruments, motion analysis devices, virtual reality simulations, and also the application of computer games is on the rise.[12-15] Currently, Objective Structured Assessment of Technical Skills (OSATS) is one of the most applied and accepted observational instruments and is considered the standard for the assessment of operative skills.[16] Originally developed at the University of Toronto by Martin et al.[17], OSATS demonstrated high reliability and validity metrics in skills laboratory settings on bench model simulations, as well as in animal models and cadaveric training models, suggesting that this instrument could measure operative skills effectively.[18,19] Subsequently, Datta et al.[20] revealed that OSATS might be of value for clinical practice by demonstrating that assessment of technical skills in an inanimate simulation model could be translated to actual surgical performance in the operating room. Since the publication of these results, OSATS

has been incorporated in numerous postgraduate medical training programs, mainly by the specialties General Surgery and Obstetrics & Gynaecology.[21-24] As part of a national initiative to reform postgraduate medical training, also the Association of Surgeons of the Netherlands introduced the OSATS instrument into its training program where it is being used to assess operative skills in the operating theater.[25]

To rely on this instrument and to defend its widespread implementation, it is imperative to have sufficient evidence comprising metrics for validity and reliability.[26,27] However, so far, only a few studies have reported on the use of OSATS in an operating theater setting.[28-32] In these mainly single center studies, evidence was either demonstrated by comparing groups with varying experience levels, ie. residents versus surgeons, performing a certain operation,[28,29] or by only evaluating surgical residents in a particular phase of a training program.[30-32] Therefore, it remains to be demonstrated whether OSATS can discriminate between surgical residents with different training levels. We conducted a multicenter study with the aim to evaluate whether OSATS is a valid and reliable instrument to assess the operative skills of surgical residents in the operating theater.

METHODS

Study design and setting

This prospective observational study was conducted in the operating rooms of 7 hospitals within a surgical training region of the Netherlands, including 1 university hospital and 6 district hospitals, from July 2012 to February 2013. In the Netherlands, training in general surgery lasts 6 years and comprises 3 consecutive training phases. The first phase consists of a 2-year training period that is shared with other surgical specialties, including cardiothoracic surgery, orthopaedics, plastic surgery and urology. Then, a second phase of training is entered that is completely focused on training in general surgery. Finally, 2 years of subspecialty training in one of the following focus areas have to be completed: gastrointestinal surgery, surgical oncology, trauma surgery, vascular surgery.[33] The objective was to assess each participating surgical resident while performing 5 different types of operations on 6 occasions, using a modified OSATS instrument. Assessments could be performed by any board certified surgeon working at the surgical department in the participating hospitals.

Participants

The program directors (n=7) of the hospitals within this training region were asked to list the names of surgical residents eligible to participate in this study, taking into account rotation schedules and the time frame of the study. From these, 24 were selected based on postgraduate training year (PGY) and invited to participate. After receiving informed consent, they were equally classified into 3 groups, each representing a different phase of the 6-year training program: group 1 (PGY1-2), group 2 (PGY3-4) and group 3 (PGY5-6).

OSATS

The original OSATS concept consists of a 3-part assessment form, including a task-specific checklist, a global rating scale, and a pass/fail judgment.[17] From this original concept, the Association of Surgeons of the Netherlands, bearing responsibility for the format and content of the surgical residency training program, adopted only the Global Rating Scale, which has shown to be superior in terms of reliability and validity compared to task-specific checklists,[17,28] in a modified version. Modifications included merging of the domains 'Knowledge of Instruments' and 'Instrument Handling', and adding of the domains 'Indication for Surgery' and 'Perioperative Management'. Furthermore, the 5-point Likert scale was replaced by a 7-point Likert scale with scores ranging from 4 to 10. In addition, 2 rating scales were added: an Overall Performance Scale and an Alphabetic Summary Scale. On the Overall Performance Scale, both surgical resident and assessor, mark a judgment of the performance shown, with awarded scores, ranging from 4 to 10, indicating whether an operation has been performed below, according to, or above expectation. As this study was aimed at evaluating surgical residents' operative skills, scores on the Overall Performance Scale assigned by surgical residents were excluded from data analysis. The Alphabetic Summary Scale provides an assessment of the ability to perform the operation on that occasion, using a 5-level alphabetic scale (level A, assists adequately; level B, competent to perform under strict supervision; level C, competent to perform under limited supervision; Level D, competent to perform without supervision; level E, supervises and educates the procedure). Finally, there is a text box available that can be used by the assessor to motivate the assigned scores on the different OSATS rating scales and to provide feedback on the specific elements of an operation needing improvement. An example of this modified OSATS instrument is shown in Figure 1.

Figure 1. Description of the OSATS instrument used by the Association of Surgeons of the Netherlands to assess surgical residents' operative skills in the operating room. Modified from Martin et al.[17]

Objective Structured Assessment of Technical Skills (OSATS)

Name: Postgraduate year of training: 1 2 3 4 5 6
Assessor: Date:
Operation:

Global Rating Scale

Domain1: Indication for Surgery						
4	5	6	7	8	9	10
No understanding of pathology		Basic understanding of pathology			Excellent understanding of pathology	
Domain 2: Respect for Tissue						
4	5	6	7	8	9	10
Frequently used unnecessary force on tissue or caused damage by inappropriate use of instrument		Careful handling of tissue but occasionally caused inadvertent damage			Consistently handled tissues appropriately with minimal damage	
Domain 3: Time and Motion						
4	5	6	7	8	9	10
Many unnecessary moves		Efficient time/motion but some unnecessary moves			Clear economy of movement and maximum efficiency	
Domain 4: Knowledge & Handling of Instruments						
4	5	6	7	8	9	10
Repeatedly awkward and unsure, inappropriate use of instruments		Occasionally stiff and awkward, mostly appropriate choice and use of instruments			Fluid moves and obviously familiar with the instruments	
Domain 5: Use of Assistants						
4	5	6	7	8	9	10
Consistently placed assistants poorly or failed to use assistants		Appropriate use of assistants most of the time			Strategically used assistants to the best advantage at all times	
Domain 6: Flow of Operation						
4	5	6	7	8	9	10
Frequently stopped operating and seemed unsure of next move		Demonstrated forward planning with reasonable progression of the operation			Obviously planned course of the operation with effortless flow from one move to the next	
Domain 7: Knowledge of Specific Procedure						
4	5	6	7	8	9	10
Deficient knowledge and needed specific instruction at most steps		Knew all important steps of the operation			Demonstrated familiarity with all aspects of the operation	
Domain 8: Perioperative Management						
4	5	6	7	8	9	10
Chaotic and incomplete		Performed well but needs instructions			Independently, carefully and complete	

Overall Performance Scale

	Below Expectation		According to Expectation			Above Expectation	
Overall judgment Assessor	4	5	6	7	8	9	10
	Dissatisfied		Neutral			Satisfied	
Overall judgment Resident	4	5	6	7	8	9	10
	Dissatisfied		Neutral			Satisfied	

Alphabetic Summary Scale

A	Competent to assist adequately
B	Competent to perform the operation under strict supervision
C	Competent to perform the operation under limited supervision
D	Competent to perform the operation unsupervised
E	Competent to supervise and educate the operation

Feedback

Operations

An overview of the selected operations with prescribed surgical technique is shown in Table 1. These operations were chosen in accordance with the program directors of the participating hospitals for the purpose of this study because they are common and regularly performed by surgical residents, reflect the broad-based character of training in general surgery, and vary in complexity and setting. For example, open and laparoscopic operations were assessed, as well as elective and emergency procedures.

Assessment instructions

Prior to the study, all surgical residents and attending surgeons of the participating hospitals were briefed on the purpose of the study and received printed instructions on how OSATS assessments should be completed. Surgical residents had to perform the operations as the primary surgeon according to their usual practice and were instructed to register an OSATS of every performed operation that was selected for this study. Surgeons were instructed to act as first assistant, allowing surgical residents to perform the operation within the limits of their competence and to intervene only in the interest of patient safety. Scoring of the modified OSATS instrument was in line with national training directives set by the Association of Surgeons of the Netherlands, prescribing that not all domains of the Global Rating Scale require completion, and that the phase of training should be taken into account when scoring the Overall Performance Scale. Finally, it was highlighted that assessments had to be completed directly after the operation and scores should not be based on previous operating theater experiences.

Table 1. Overview of the selected operations with prescribed surgical technique.

Indication for Surgery	Surgical Technique
Symptomatic cholecystolithiasis	Laparoscopic cholecystectomy (4-port 2-handed technique)
Inguinal hernia	Lichtenstein repair
Inguinal hernia	Totally extraperitoneal procedure (TEP)
Proximal femoral fracture	Internal fixation by cannulated hip screws, sliding hip screw or intramedullary nailing
Breast tumor	Lumpectomy or mastectomy (± followed by sentinel node biopsy or axillary lymph node dissection)

Statistical analysis

Statistical analyses were performed using IBM SPSS version 20.0 (IBM Corp., Armonk, NY, US). For all performance scores on the different rating scales of the modified OSATS instrument, mean values and standard deviations were calculated. As not all domains of the Global Rating Scale require completion during an assessment, scores were based on the average score of the completed domains. To perform statistical analyses, scores on the Alphabetic Summary Scale were interpreted as continuous data and are displayed numerically (score A=1, score E=5). As our data were not normally distributed, nonparametric tests were used for reliability and validity analyses. Construct validity, defined as the ability of an instrument to differentiate between groups with varying experience levels, was determined by comparing scores on the different rating scales across the 3 groups of surgical residents, using the Kruskal-Wallis test. The Mann-Whitney test was applied for post-hoc analyses. Cronbach's coefficient was determined to establish the internal consistency reliability of the Global Rating Scale. The value of this coefficient reflects to which extent different domains contribute to the overall score on this particular scale. Spearman's correlation coefficients were calculated to determine correlations between the different rating scales. For all tests, $p<0.05$ was considered statistically significant different.

RESULTS

Of the 24 recruited surgical residents, 6 dropped out before the study started. One resident of group 1 (PGY 1-2) left the training program, whereas 5 residents of group 3 (PGY 5-6) withdrew and indicated they were not able to participate due to the increased complexity of rotation schedules resulting from subspecialty training. The 18 surgical residents performed 255 operations that were completed with an OSATS assessment, revealing a discrepancy between the total number of 540 operations (18 residents * 5 operations * 6 occasions) expected to be performed within the time frame of this study and the actual number of procedures carried out. Of these 255 operations, 6 were excluded because of peroperatively new findings (cholecystitis, n=2) or due to applied surgical technique (hip hemi-arthroplasty, n=2; axillary lymph node dissection without breast surgery, n=2). The remaining 249 operations included 80 laparoscopic cholecystectomies, 69 inguinal hernia repairs, of which 45 were performed by a Lichtenstein repair and 24 by a totally extraperitoneal procedure (TEP), 41 proximal femoral fracture osteosyntheses and 59 breast tumor operations (in some cases followed by sentinel lymph node biopsy or axillary lymph node dissection). The operations were evaluated by 41 different surgeons who completed a mean number of 6 (median 4, range 1-21) OSATS assessments per person. The mean number of operations performed per surgical resident was 14 (median 13, range 2-27), with each surgical resident having a mean number of 5 (median 4, range 2-8) different assessors completing the OSATS assessments.

Construct validity

For each of the operations, performance scores on the different rating scales of the modified OSATS instrument are presented in Table 2. Comparisons of these scores are displayed in Figures 2a-e. For the laparoscopic cholecystectomy (Fig. 2a), scores on all rating scales improved across the different groups ($p\leq0.005$), except for scores on the Overall Performance Scale between group 2 and group 3 ($p=0.057$). Scores for the Lichtenstein repair (Fig. 2b) demonstrated differences on all rating scales between group 1 and group 2 ($p<0.001$). In contrast, no statistically significant differences were found between group 2 and group 3. For the TEP (Fig. 2c), differences were present in the scores on the Overall Performance Scale between group 1 and group 2 ($p=0.036$), and in the scores on the Global Rating Scale ($p=0.003$) and the Alphabetic Summary Scale ($p=0.001$) between group 2 and group 3. Analysis of the scores for proximal femoral fracture osteosyntheses (Fig. 2d) revealed only a difference on the Alphabetic Summary Scale between group 1 and group 2 ($p=0.011$). No significant differences were found in the scores on any of the other rating scales across the different groups.

Table 2. Performance scores on the different rating scales of the modified OSATS instrument for the 3 groups of surgical residents.

		Group 1 (PGY 1-2) n=7 *		Group 2 (PGY 3-4) n=8 *		Group 3 (PGY 5-6) n=3 *	
	Rating Scale	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**
Laparoscopic cholecystectomy		29 §		35 §		16 §	
	GRS (=0.93)#	7.5	0.77	8.4	0.58	8.8	0.49
	OPS	7.7	0.91	8.3	0.83	8.8	0.45
	ASS	2.7	0.55	3.9	0.68	4.8	0.40
Lichtenstein repair		23 §		13 §		9 §	
	GRS (=0.94)#	7.8	0.54	8.7	0.26	8.8	0.63
	OPS	7.8	0.66	8.9	0.29	8.7	0.50
	ASS	2.5	0.51	4.6	0.51	5.0	0.00
Totally extraperitoneal procedure (TEP)		10 §		6 §		8 §	
	GRS (=0.93)#	7.3	0.35	7.7	0.59	9.0	0.53
	OPS	7.0	0.00	7.7	0.52	8.4	1.06
	ASS	2.6	0.53	2.7	0.52	4.8	0.46
Proximal femoral fracture osteosynthesis		18 §		11 §		12 §	
	GRS (=0.94)#	7.8	0.63	8.3	0.71	8.5	0.49
	OPS	8.1	0.88	8.4	0.92	8.6	0.52
	ASS	2.9	0.66	4.0	0.95	4.7	0.65
Breast tumor surgery		19 §		28 §		12 §	
	GRS (=0.95)#	7.2	0.73	8.1	0.66	8.7	0.49
	OPS	7.1	0.64	8.1	0.69	8.1	0.79
	ASS	2.2	0.44	3.2	0.70	4.6	0.52

*number of surgical residents per group, §number of operations performed per group, # internal consistency reliability of the GRS

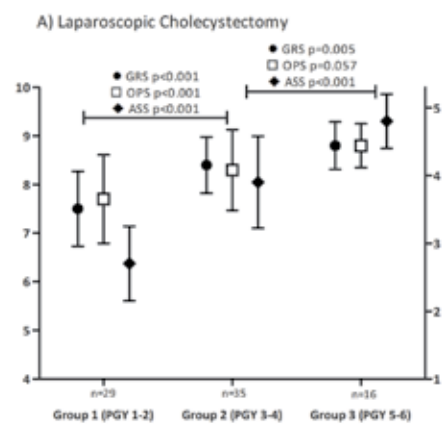
PGY (postgraduate training year); SD (standard deviation); GRS (global rating scale) and OPS (overall performance scale) with scores ranging from 4 to 10; ASS (alphabetic summary scale) with scores ranging from 1 [assists adequately] to 5 [supervises and educates]

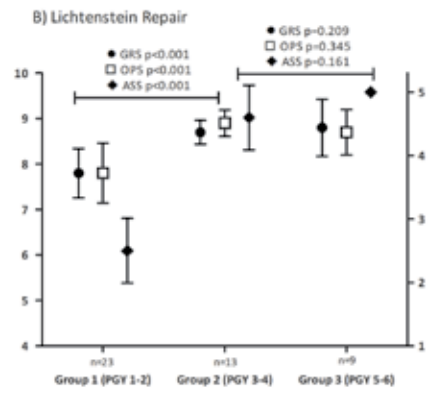
Scores for the breast tumor operations (Fig. 2e) increased substantially on all rating scales ($p \leq 0.010$), except for scores on the Overall Performance Scale between group 2 and group 3 ($p=0.965$).

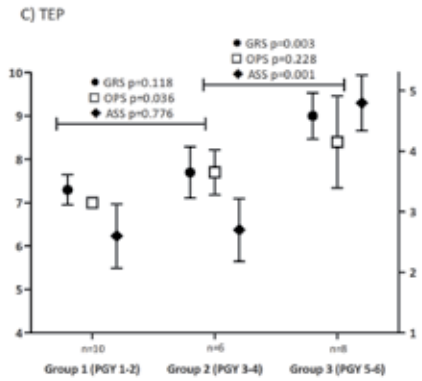
Internal consistency reliability

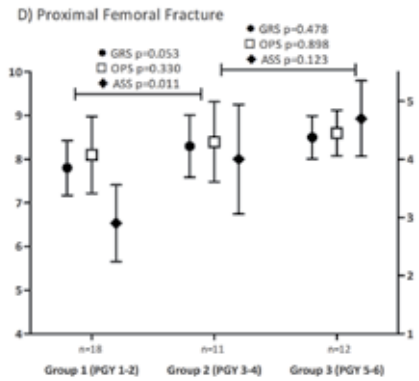
The internal consistency reliability of the Global Rating Scale for each of the operations was excellent, with Cronbach's α coefficients ranging from 0.93 to 0.95 (Table 2), indicating that scores on all domains were correlated with the overall score on this particular rating scale.

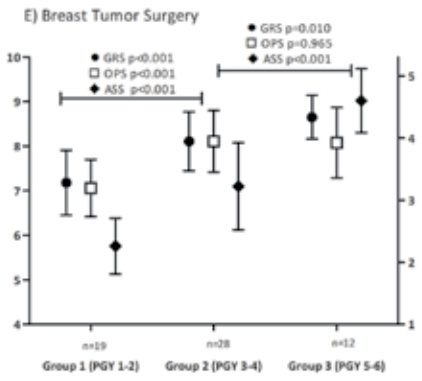
Figure 2a-e. Performance scores on the different rating scales of the modified OSATS instrument for each of the operations: **a)** laparoscopic cholecystectomy, **b)** Lichtenstein repair, **c)** totally extraperitoneal procedure (TEP), **d)** proximal femoral fracture osteosynthesis, **e)** breast tumor surgery.











Data are presented as means and standard deviations (exact values are depicted in Table 2). Results are based on 18 surgical residents: group 1 (PGY1-2, n=7); group 2 (PGY3-4, n=8); group 3 (PGY5-6, n=3). Left y-axis: GRS (Global Rating Scale) and OPS (Overall Performance Scale) with scores ranging from 4 to 10; Right y-axis ASS (Alphabetic Summary Scale) with scores ranging from 1 to 5. Comparisons between groups, as an indication for construct validity, were made using the Kruskal-Wallis test. Post hoc analyses were performed with the Mann-Whitney test.

Correlations

Scores on the Global Rating Scale correlated strongly with scores on the Overall Performance Scale (r=0.76, p<0.001) and scores on the Alphabetic Summary Scale (r=0.75, p<0.001). A weaker correlation was observed between scores on the Overall Performance Scale and scores on the Alphabetic Summary Scale (r=0.62, p<0.001). In Table 3, correlation coefficients between scores for the individual domains of the Global Rating Scale and the scores on both the Overall Performance Scale and Alphabetic Summary Scale are shown. All correlation coefficients were statistically significant, with all p-values <0.001. The domain 'Flow of Operation' (r=0.78) demonstrated the strongest correlation with scores on the Overall Performance Scale, whereas the domains 'Indication for Surgery' (r=0.45) and 'Perioperative Management' (r=0.53) showed the weakest correlations. The domains 'Time and Motion', 'Knowledge and Handling of Instruments' and 'Use of Assistants' demonstrated the strongest correlations with scores on the Alphabetic Summary Scale. Again, the weakest correlations were found for the domains 'Indication for Surgery' (r=0.34) and 'Perioperative Management' (r=0.39).

Table 3. Spearman's correlation coefficients between scores for the individual domains of the global rating scale (GRS) and scores on the overall performance scale (OPS) and the alphabetic summary scale (ASS).

	OPS		ASS	
Domains of the GRS	**n**	**Spearman's rho**	**n**	**Spearman's rho**
Indication for Surgery	183	0.45	184	0.34
Respect for Tissue	232	0.66	234	0.61
Time and Motion	231	0.70	233	0.74
Knowledge and Handling of Instruments	226	0.68	228	0.74
Use of Assistants	215	0.64	217	0.74
Flow of Operation	232	0.78	234	0.69
Knowledge of Specific Procedure	232	0.66	233	0.61
Perioperative Management	210	0.53	211	0.39

Note: All correlations were statistically significant with all p-values <0.001

DISCUSSION

The aim of this study was to examine whether OSATS is a valid and reliable instrument to assess the operative skills of surgical residents in the operating theater. Compared to previous studies reporting on the evaluation of the OSATS instrument in an operating theater environment,[20,28-32] a study design in more resemblance with daily clinical practice was chosen. The participants in this study were selected from all postgraduate training years within a regionally organized surgical training program and classified into 3 groups, with each group representing a particular phase of the training program (group 1 PGY 1-2; group 2 PGY 3-4; group 3 PGY 5-6). In this way, classification based on a vast difference in operative experience between the studied groups, such as in previous studies, was avoided.[28,29] In addition, it also allowed to study surgical residents of various training levels compared to other previous studies examining the use of the OSATS instrument in the operating room in which merely residents from certain postgraduate training years were enrolled.[30-32] Furthermore, a variety of operations were evaluated in comparison to other studies in which only a specific procedure was assessed, such as laparoscopic cholecystectomy,[28,30] or carotid endarterectomy.[29] This approach allowed to evaluate the validity and reliability of the OSATS instrument to assess operations with varying difficulty levels commonly performed during training in general surgery.

Construct validity was determined by examining whether performance scores on the different scales of the modified OSATS instrument increased along with the training level of the surgical residents. Comparisons of these scores between the groups revealed mixed construct validity across the various rating scales, depending on the difficulty level of the operation performed. Scores for the laparoscopic cholecystectomy (Fig. 2a), for example, showed significant differences on almost all rating scales between the 3 groups. This finding reflects that laparoscopic surgery requires optimal video-eye-hand coordination, a skill gradually mastered during surgical residency training. For the Lichtenstein repair (Fig. 2b), which is one of the first and relatively simple operations residents are exposed to, significant differences were, as expected, only found between group 1 and group 2. In contrast, for the TEP (Fig. 2c) significant differences were mainly observed between group 2 and group 3, demonstrating that this operation is complex and requires advanced laparoscopic skills which are being mastered in a later phase of the training program. Comparing scores for the proximal femoral fractures osteosyntheses (Fig. 2d) revealed almost no significant differences across the different groups of residents. A potential explanation for this lack of construct validity could be the simplicity of this operation. Comparisons of the scores for the breast tumor operations (Fig. 2e) showed significant differences on almost all rating scales between the 3 groups. However, this evidence for construct validity might be biased by the varying difficulty level of this operation as lumpectomy or mastectomy occasionally was followed by sentinel node biopsy or axillary lymph node dissection. These findings generally indicate that this modified OSATS instrument has the potential to establish learning curves, allowing adequate monitoring of surgical residents' progress in achieving operative competence.

A more precise examination of the different rating scales of this modified OSATS instrument revealed several interesting findings. First, it was found that the Overall Performance Scale demonstrated the least significant differences, indicating that this rating scale had the weakest discriminative ability. As scores on this rating scale display whether a surgical resident has performed an operation according to the expectation of the assessor completing the assessment, we conclude that scores on this scale do not provide an objective indication for operative competence, but rather reflect a subjective perception of surgical performance that might be tainted by interpersonal relationships. Therefore, use of the Overall Performance Scale should be reconsidered, especially because this rating scale is not part of the original OSATS concept. The Alphabetic Summary Scale, on the other hand, revealed the most significant differences between the different groups, indicating that this scale had the greatest discriminative power. This is an important finding that demonstrates the additional value of the Alphabetic Summary Scale and supports its application for use in clinical practice. In our opinion, this rating scale also provides the most objective indication for operative skill because scores reflect whether a surgical resident is considered competent to perform an operation or not. Such a clear definition of competence is much more informative than quantifying surgical skills by circling scores on numerical rating scales.

When correlation statistics were calculated to evaluate the relationship between scores for the individual domains of the Global Rating Scale and scores on both the Overall Performance Scale and Alphabetic Summary Scale (Table 3), strongest correlations were found for the domains related to aspects of operative skill, such as 'Respect for Tissue', 'Time and Motion', 'Use of Assistants' etc. In contrast, weakest correlations were observed for the domains 'Indication for Surgery' and 'Perioperative Management'. Interestingly, it were also these 2 domains that were added to the original concept of the Global Rating Scale. Because scores on these domains, although critical components of operative care, do not reflect operative skill in any way, we recommend that both domains should be deleted from the modified version of the Global Rating Scale used in this study and should not be adopted by others.

The reliability of this modified OSATS instrument was determined by calculating the internal consistency reliability of the Global Rating Scale, which was excellent for all 5 operations included in this study. Other reliability metrics that are frequently determined when an assessment tool is evaluated are inter-test reliability and inter-rater reliability. Inter-test reliability reflects the ability of an instrument to generate similar results when a specific operation is evaluated by the same rater on different time points. Inter-rater reliability refers to the extent of agreement between scores when multiple raters simultaneously evaluate the same operation.[26,27] Ideally, to avoid bias when determining these reliability metrics, consecutive operations should be evaluated by multiple raters who are blinded to the training level of the surgical residents being assessed. However, defining a large group of surgical residents who are able to perform a high number of consecutive operations within a particular time frame could not be realized due to stringent work hour restrictions and subspecialty training. Moreover, finding attending surgeons who want to participate in such a study will pose an even greater challenge, especially because they have to originate from other hospitals if they need to act as blinded raters. These requirements illustrate that conducting high-quality studies, with the objective to assess operative skills in an uncontrolled setting like an operating room, is a challenging but above all extremely difficult undertaking. The loss of surgical residents in the senior group (PGY 5-6), as well as the considerable discrepancy between the expected number of 540 operations and the actual number of 255 operations that were performed within the time frame of this study exemplify this difficulty.

The high values and limited spread of the scores on both the Global Rating Scale and Overall Performance Scale, with means ranging from 7.0 to 9.0 (on a 4-10 rating scale), were considered a remarkable finding. A possible explanation for this finding could be the relative simplicity of the selected operations. However, similar findings were also observed in other studies examining the operative skills of surgical residents in an operating theater setting, using observational instruments.[28,31,34] Consequently, one might wonder whether high performance scores accurately reflect operative competence. In this context, Williams et al.[35] investigated the prognostic value of clinical performance ratings within a general surgical residency program in the United States. They demonstrated that a vast majority of the residents received 'Excellent' or 'Very Good' scores and only a small minority received 'Fair' or 'Poor' scores, from which the authors concluded that such ratings have no directly interpretable meaning. This finding might have important consequences as we move forward into a competency-based assessment model of surgical residency training. First, high scores make it difficult, if not impossible, to identify surgical residents who are incompetent and function substandardly compared to their peers. Secondly, high scores do not provide any information on the specific elements of an operation that can be improved. In addition, there are

several other issues associated with the evaluation of operative skills in an operating theater environment that potentially could cause bias. As surgical residents and attending surgeons closely collaborate in daily practice, it is not unlikely that personal preferences may influence scoring behavior and thus assessors might be chosen for other reasons than their ability to complete an assessment. Finally, it is almost impossible to perform a standardized assessment for a specific operation due to anatomic variations and varying patient characteristics. These issues open the debate regarding the role OSATS assessments should fulfill in clinical practice; either formative or summative. Formative assessments allow surgical residents and their trainers to monitor progression and can help structuring learning processes by providing constructive feedback. Summative assessments, on the other hand, have the potential to be used for credentialing or certification purposes. We believe, however, that such important decisions should be based on instruments that provide a valid and reliable indication for operative competence. In a systematic review addressing the validity and reliability of various methods for surgical skills assessment, it was recommended that the OSATS instrument should not be used for summative evaluation purposes due to a lack of high-level evidence.[16] This is an important message that should be taken into account by responsible professional associations and other stakeholders involved in postgraduate medical training who also have adopted the OSATS instrument for surgical skills assessment.

CONCLUSION

This is the first multicenter study that prospectively has evaluated whether OSATS is a valid and reliable instrument to assess the operative skills of surgical residents enrolled in various phases of a training program in general surgery, while performing a variety of operations in an operating theater setting. Despite that a modified version of the original OSATS concept was used, limiting comparison of our findings to those in other studies, this study has demonstrated that performance scores on the different rating scales improved along with the training level of the surgical residents, which constitutes evidence for construct validity. Evidence for reliability was shown by the high internal consistency reliability coefficients for scores on the Global Rating Scale. Given its weak discriminative ability, use of the Overall Performance Scale should be reconsidered. The Alphabetic Summary Scale, on the other hand, seems to be of additional value for clinical practice. Nevertheless, it must be pointed out that, despite observational instruments such as OSATS, may currently be the most accepted and easily available assessment method to evaluate surgical residents' operative skills in the operating room, performance scores on these instruments are inherently associated with subjectivity and do not necessarily provide an accurate and objective indication for operative competence.

REFERENCES

1. Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. BMJ 2003;327(7422): 1032-1037.
2. Reznick RK. Teaching and testing technical skills. Am J Surg 1993;165(3): 358-361.
3. Glomsaker TB, Soreide K. Surgical training and working time restriction. Br J Surg 2009;96(4): 329-330.
4. Kairys JC, McGuire K, Crawford AG, Yeo CJ. Cumulative operative experience is decreasing during general surgery residency: a worrisome trend for surgical trainees? J Am Coll Surg 2008;206(5): 804-811; discussion 811-803.
5. Curry L, Wergin JF. Educating professionals: Responding to new expectations for competence and accountability. Jossey-Bass: San Francisco, CA, 1993; 379.
6. Brennan MF, Debas HT. Surgical education in the United States: portents for change. Ann Surg 2004;240(4): 565-572.
7. Collins JP, Civil ID, Sugrue M, Balogh Z, Chehade MJ. Surgical education and training in Australia and New Zealand. World Journal of Surgery 2008;32(10): 2138-2144.
8. Frank JR, Langer B. Collaboration, communication, management, and advocacy: teaching surgeons new skills through the CanMEDS Project. World J Surg 2003;27(8): 972-978; discussion 978.
9. Smith AJ, Aggarwal R, Warren OJ, Paraskeva P. Surgical training and certification in the United kingdom. World J Surg 2009;33(2): 174-179.
10. Epstein RM. Assessment in medical education. N Engl J Med 2007;356(4): 387-396.
11. Baldwin PJ, Paisley AM, Brown SP. Consultant surgeons' opinion of the skills required of basic surgical trainees. Br J Surg 1999;86(8): 1078-1082.
12. Ahmed K, Miskovic D, Darzi A, Athanasiou T, Hanna GB. Observational tools for assessment of procedural skills: a systematic review. Am J Surg 2011;202(4): 469-480 e466.
13. 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.
14. Aggarwal R, Grantcharov TP, Eriksen JR, Blirup D, Kristiansen VB, Funch-Jensen P, et al. An evidence-based virtual reality training program for novice laparoscopic surgeons. Ann Surg 2006;244(2): 310-314.
15. Graafland M, Schraagen JM, Schijven MP. Systematic review of serious games for medical education and surgical skills training. Br J Surg 2012;99(10): 1322-1330.
16. van Hove PD, Tuijthof GJ, Verdaasdonk EG, Stassen LP, Dankelman J. Objective assessment of technical surgical skills. Br J Surg 2010;97(7): 972-987.
17. 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.

18. 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.
19. Anastakis DJ, Regehr G, Reznick RK, Cusimano M, Murnaghan J, Brown M, et al. Assessment of technical skills transfer from the bench training model to the human model. Am J Surg 1999;177(2): 167-170.
20. Datta V, Bann S, Beard J, Mandalia M, Darzi A. Comparison of bench test evaluations of surgical skill with live operating performance assessments. J Am Coll Surg 2004;199(4): 603-606.
21. Boehler ML, Schwind CJ, Rogers DA, Ketchum J, O'Sullivan E, Mayforth R, et al. A theory-based curriculum for enhancing surgical skillfulness. J Am Coll Surg 2007;205(3): 492-497.
22. Chipman JG, Schmitz CC. Using objective structured assessment of technical skills to evaluate a basic skills simulation curriculum for first-year surgical residents. J Am Coll Surg 2009;209(3): 364-370 e362.
23. 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. Am J Obstet Gynecol 2007;197(5): 544 e541-546.
24. Swift SE, Carter JF. Institution and validation of an observed structured assessment of technical skills (OSATS) for obstetrics and gynecology residents and faculty. Am J Obstet Gynecol 2006;195(2): 617-621; discussion 621-613.
25. Hopmans CJ, den Hoed PT, Wallenburg I, van der Laan L, van der Harst E, van der Elst M, et al. Surgeons' attitude toward a competency-based training and assessment program: results of a multicenter survey. J Surg Educ 2013;70(5): 647-654.
26. 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.
27. Oropesa I, Sanchez-Gonzalez P, Lamata P, Chmarra MK, Pagador JB, Sanchez-Margallo JA, et al. Methods and tools for objective assessment of psychomotor skills in laparoscopic surgery. J Surg Res 2011;171(1): e81-95.
28. 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.
29. Beard JD, Choksy S, Khan S. Assessment of operative competence during carotid endarterectomy. Br J Surg 2007;94(6): 726-730.
30. 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.
31. Hiemstra E, Kolkman W, Wolterbeek R, Trimbos B, Jansen FW. Value of an objective assessment tool in the operating room. Can J Surg 2011;54(2): 116-122.
32. Niitsu H, Hirabayashi N, Yoshimitsu M, Mimura T, Taomoto J, Sugiyama Y, et al. Using the Objective Structured Assessment of Technical Skills (OSATS) global rating scale to evaluate the skills of surgical trainees in the operating room. Surg Today 2013;43(3): 271-275.

33. Borel-Rinkes IH, Gouma DJ, Hamming JF. Surgical training in the Netherlands. World J Surg 2008;32(10): 2172-2177.
34. 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.
35. Williams RG, Dunnington GL. Prognostic value of resident clinical performance ratings. J Am Coll Surg 2004;199(4): 620-627.

Chapter 5

Impact of the European Working Time Directive (EWTD) on the operative experience of surgical residents

Cornelis J. Hopmans
Pieter T. den Hoed
Lijckle van der Laan
Erwin van der Harst
Maarten van der Elst
Guido H.H. Mannaerts
Imro Dawson
Reinier Timman
Bas P.L. Wijnhoven
Jan N.M. IJzermans

Adapted from: Surgery. 2015 Apr;157(4):634-41.

ABSTRACT

Background: In Europe and the United States, work hour restrictions are considered to be particularly burdensome for residents in surgical specialties. The aim of this study was to examine whether the progressive reduction of the working week to 48 hours resulting from implementation of the European Working Time Directive (EWTD) has affected the operative experience of surgical residents.

Methods: This study was conducted in a surgical training region in the Netherlands, consisting of 1 university hospital and 6 district hospitals. Operating records summarizing the surgical procedures performed as 'primary surgeon' in the operating theater for different grades of surgeons were retrospectively analyzed for the period 2005-2012 by the use of linear regression models. Surgical procedures performed by residents were considered the main outcome measure.

Results: In total, 235,357 surgical procedures were performed, including 47,458 (20.2%) in the university hospital and 187,899 (79.8%) in the district training hospitals (n=5). For residents in the university hospital, the mean number of surgical procedures performed per 1.0 full-time equivalent increased from 128 operations in 2005 to 204 operations in 2012 (p=0.001), whereas for residents in district training hospitals no significant differences were found over time. The mean (±SD) operative caseload of 64 residents who completed the 6-year training program between 2005-2012 was 1,391 ± 226 (range: 768-1,856). A comparison of the operative caseload according to year of board certification showed no significant difference.

Conclusion: Implementation of the EWTD has not adversely affected the number of surgical procedures performed by residents within a surgical training region in the Netherlands.

INTRODUCTION

Within the past decade, work hour restrictions have been enforced by law in most Western countries. In the European Union, implementation of the European Working Time Directive (EWTD - 2003/88/EC) has progressively limited the average working week for medical residents from a maximum of 58 hours in August 2004, to 56 hours from August 2007, and to ultimately 48 hours from August 2009. In addition to this substantial reduction in working hours, strict requirements for daily break moments, weekly rest periods and regulation of night and shift-work are stipulated, further decreasing the amount of training time and availability during daytime.[1,2] In a similar fashion, in the United States (US), the Accreditation Council for Graduate Medical Education (ACGME) instituted an 80-hour limited working week in 2003 and recently, as recommended in a report by the Institute of Medicine in 2008, stricter regulations have been implemented in 2011.[3]

An important aim of the work hour restrictions is to protect the health and safety of medical residents. In a review, conducted by Curet et al.,[4] it was concluded that the quality of life of US residents has improved significantly since the introduction of an 80-hour working week, as evidenced by lower depression scores,[5] a decrease in motor vehicle accidents after working extended shifts,[6] and an increased amount of sleep.[7] This improvement in quality of life has been confirmed in 2 systematic reviews concluding that the introduction of work hour restrictions has had positive effect on US residents' personal well-being and lifestyle.[8,9] Paradoxically, recently published survey studies among surgical residents in the US,[10,11] as well as European questionnaire studies conducted among residents from various surgical specialties,[12,13] revealed that work hour regulations are believed to have deteriorated the worklife balance due to an increased amount of shift-work, which is perceived as being more tiring and disruptive than a 'normal' working week.

Another aim of the work hour restrictions is to reduce the risk to patient safety posed by exhaustion and fatigue resulting from working excessive long hours. This theory is supported by evidence that elimination of extended shift-work is associated with a decrease in the number of preventable medical errors and adverse events.[14,15] In contrast, others claim that reduced working hours are associated with an increase in irregular hours of shift-work, complex rotation schedules and more frequent handovers, resulting in a loss of continuity of care that eventually will lead to impaired patient outcomes, and thus ultimately adversely will affect patient safety.[16,17]

Within the medical profession, work hour restrictions are in general being condemned for having a detrimental effect on the quality of training, as described in a large amount of literature from both sides of the Atlantic.[17-20] In particular surgical specialties, with frequent emergency cases and after-hour demands, have expressed considerable concerns and remain opposed to the restrictions.[21-23] This opposition seems primarily as a result of the perception that a decrease in surgical residents' working hours comes at the expense of both sufficient operative and clinical experience, potentially threatening the high standard of surgical care and safety of patients. These contentions also have been acknowledged in a position paper by the Section of Surgery of the European Union of Medical Specialists (UEMS), stating that 'the EWTD is in direct and severe conflict with European legislation to train competent surgical specialists'.[24] In the context of these contentions, a systematic review examining the impact of work hour restrictions on objective measurable educational outcomes in both the US and the European Union concluded that postgraduate training in the US had not been adversely affected, but that high-quality studies investigating the impact of the EWTD were lacking.[25]

In the Netherlands, significant changes have affected the structure and format of surgical residency training within the past decade. In 2009, in response to contemporary healthcare needs, a competency-based training program has been implemented. At the same time, due to pressure on health care budgets, the duration of surgical training, often lasting up to 8 years due to postresidency fellowship training, has been shortened to 6-year program in which subspecialty training is mandatory and offered during the last 2 years. On top of this significant reduction in training time, working weeks have been limited by implementation of the EWTD. The Netherlands, however, received a temporary extension to meet the 48-hour target set for August 2009, with working weeks being restricted to 52 hours in the transition period lasting until August 2011.[26,27] Although the significance of operative competence alone is of limited importance for developing proficiency during surgical training, evidence is needed to inform us about the impact of work hour restrictions on the training and education of the next generation of surgeons. The aim of this study was to evaluate whether the EWTD has had an effect on the operative experience of surgical residents.

METHODS

Setting

In the Netherlands, training in general surgery is organized in 8 training regions, each consisting of 1 university hospital and several affiliated district hospitals. To become a board certified surgeon, surgical residents have to complete a 6-year training program, of which 2 years are spent in the university hospital and the remaining 4 years in one of the district hospitals. The first 2 years of the training program, also called common trunk, are shared with other surgical specialties, including cardiothoracic surgery, orthopaedic surgery, plastic surgery and urology.

Design

This study was conducted in a surgical training region in the Netherlands, consisting of 1 university hospital and 6 district hospitals. Operating records, derived from electronic hospitals systems in which all surgical procedures performed on a yearly basis for different grades of surgeons are registered, were retrospectively analyzed for the period 2005-2012. In these records, only surgical procedures performed in the operating theater are included. Consequently, minor surgical procedures, such as simple skin excisions that are regularly performed at the outpatient clinic, were excluded. Coding of the surgical procedures is performed by the non-scrubbed operating room staff, who, in close consultation with the attending surgeon present in the operation theater, register a procedure in the name of the 'primary surgeon', defined as the individual who carried out the critical steps during a procedure. This information is saved on a personal digital account, allowing to accurately figure out who was the surgeon performing a particular operation and who acted as his/her assistant(s). Because 1 of the district hospitals joined this training region only in 2011, data from this hospital were excluded. Permission to use the data was obtained from the program directors of the participating hospitals.

Data analysis

First, changes in the number of surgical procedures performed by different grades of surgeons were analyzed, including attending surgeons (independently practicing surgical specialists), fellow surgeons (board certified surgeons pursuing subspecialty training after completion of residency training), surgical residents (physicians in training to become a surgeon) and surgical house officers (physicians who do not hold a training position). None of these definitions changed over time. Subsequently, the surgical procedures carried out by the

group of residents were further analyzed by adjusting for changes in the workload, mainly being caused by an increased desire to work part-time, as well as maternity leave.[28] To correct for these compromised work schedules and periods of absence, the departments of human resources were consulted to obtain data on the number of full-time equivalents (FTEs) being employed as surgical residents during the period studied. Thus, the number of surgical procedures performed per individual resident, i.e. per 1.0 FTE, could be determined more accurately and allows firmer conclusions to be drawn in comparison to reporting the number of operations carried out per person. Finally, the operative caseload of residents, who successfully completed the 6-year training program and obtained their board certification between 2005-2012, was examined and compared over time. As surgical procedures performed in a university hospital may differ in complexity compared to those performed in district hospitals, data for both types of hospitals were analyzed separately.

Statistical analysis

Statistical analyses were performed using IBM SPSS version 21.0 (IBM corp., Armonk, NY, US). Descriptive statistics were calculated, including percentages, means, standard deviations (SD), and ranges. Changes in the number of surgical procedures performed by the different grades of surgeons were analyzed by the use of linear regression models. The surgical house officers were used as reference group and the groups of surgical residents, fellow surgeons and attending surgeons were entered as dummified covariates. Time, and where appropriate, quadratic time, and interactions with the group dummies were entered as covariates. Similar regression statistics were used to evaluate the change in number of surgical procedures carried out per 1.0 FTE surgical resident and to compare the operative caseload of residents who obtained their board certification over the studied time frame. The level for statistical significance was set at $p<0.05$.

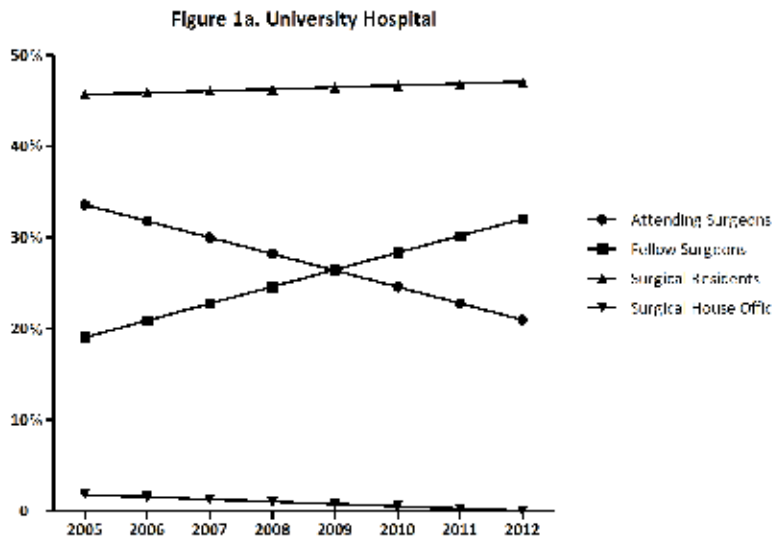
RESULTS

Within this surgical training region, a total of 235,357 surgical procedures were performed between 2005 and 2012, including 47,458 (20.2%) in the university hospital and 187,899 (79.8%) in the 5 different district hospitals (range: 29,130-50,103). In the university hospital, the total number of operations performed decreased from 6,284 surgical procedures in 2005 to 5,588 in 2012, representing an annual decline of 1.6% ($p=0.047$). In contrast, the total number of operations carried out in the district hospitals increased by 2.3% on a yearly basis from 21,743 surgical procedures in 2005 to 25,240 surgical procedures in 2012 ($p=0.001$). To correct for these significant institutional volume differences, the change in operative caseload for the different grades of surgeons are presented as proportions.

In the university hospital, the percentage of operations carried out by attending surgeons decreased from 33.6% in 2005 to 20.9% in 2012 ($p=0.036$). In contrast, for the fellow surgeons an increase was found from 19.0% in 2005 to 32.0% in 2012 ($p=0.006$). The percentage of operations performed by surgical residents ($p=0.515$) and surgical house officers ($p=0.608$) did not change significantly over time (Fig. 1a). In the district hospitals, the percentage of operations carried out by attending surgeons decreased from 41.8% in 2005, to a minimum of 37.9% in 2008, followed by an increase to 48.5% in 2012, representing a significant variance over time ($p=0.004$). The opposite was found for the surgical residents, for whom the percentage of performed operations increased from 55.0% in 2005 to a maximum of 58.7% in 2008 and then declined to 45.3% in 2012 ($p=0.001$). Over the period studied, no differences were found for both the groups of fellow surgeons ($p=0.764$) and surgical house officers ($p=0.870$). These results can be seen in Figure 1b.

Subsequently, the surgical procedures carried out by the surgical residents were further analyzed by adjusting for changes in the workload. In the university hospital, the mean number of operations per 1.0 FTE increased substantially from 128 surgical procedures in 2005 to 204 surgical procedures in 2012 ($p=0.001$), while the number of FTEs showed a decrease, from 21.5 FTE in 2005 to 12.1 FTE in 2012 ($p=0.001$). These results can be seen in Figure 2a. As shown in Figure 2b, for residents in the district hospitals, the mean number of operations per 1.0 FTE did not change significantly ($p=0.607$), and also the number of FTEs was stable over time ($p=0.812$).

Figure 1a-b. Proportional change in the number of surgical procedures performed by different grades of surgeons in the university hospital (Fig. 1a) and the district hospitals (Fig. 1b) between 2005 and 2012. Implementation of the EWTD in the Netherlands: 2004 max 58 h; 2007 max 56 h; 2009 max 52 h; 2011 max 48 h.



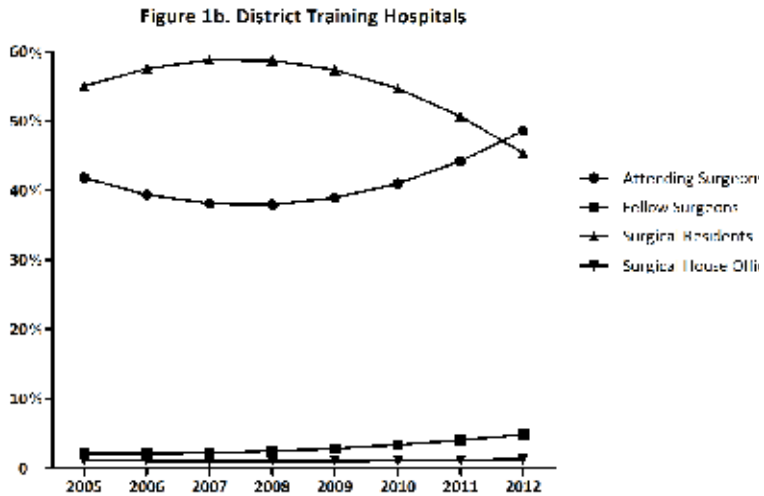
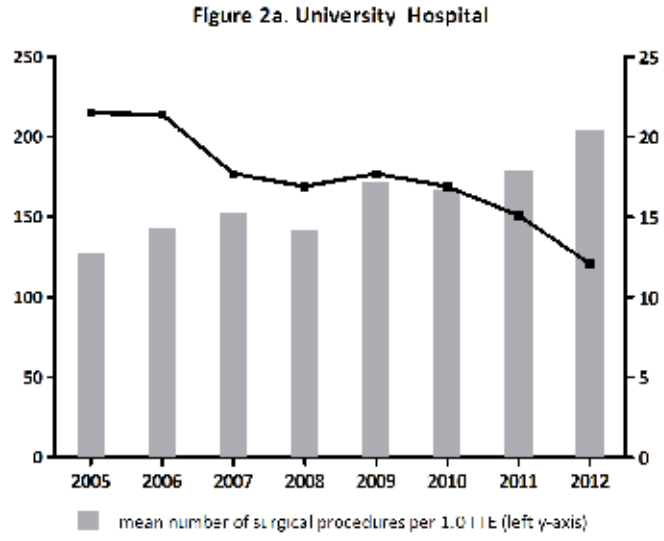
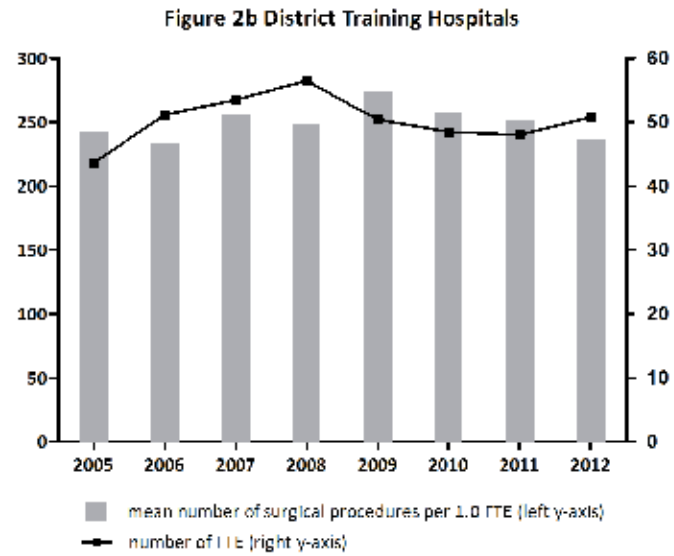


Figure 2a-b. Mean number of surgical procedures performed per 1.0 full-time equivalent (FTE) surgical resident in the university hospital (Fig. 2a) and the district hospitals (Fig. 2b) between 2005 and 2012. Implementation of the EWTD in the Netherlands: 2004 max 58 h; 2007 max 56 h; 2009 max 52 h; 2011 max 48 h.





Finally, the operative caseload of surgical residents on completion of the 6-year training program was examined. Within this training region, 78 residents were enrolled in the surgical training program between 2005 and 2012, of which 3 stopped their training in an early stage. For the data analysis of the remaining 75, 11 had to be excluded, either because a part of the training program was completed in another training region and therefore data were missing (n=7), or due to incompleteness of the data supplied (n=4). For the remaining 64, the mean (±SD) operative caseload was 1,391 ± 226 (range: 768-1,856). In Table 1, an overview of these numbers and the mean operative caseload according to year of board certification is shown. A comparison of these means revealed no difference (p= 0.129), indicating that the operative caseload surgical residents are exposed to during their training has been stable over the period studied, despite an increasing number completing the training program.

Table 1. Operative caseload of surgical residents on completion of the 6-year training program according to year of board certification for the period 2005-2012.

			Excluded from data analysis		
Year	**Board certified surgeons**	**Surgical procedures performed, mean ± SD**	**Stop training**	**Mutation of training region**	**Missing data**
2005	3	1,484 ± 218	-	1	1
2006	7	1,329 ± 194	-	1	2
2007	7	1,490 ± 297	-	-	1
2008	9	1,430 ± 147	-	-	-
2009	5	1,310 ± 237	1	2	-
2010	11	1,291 ± 158	-	2	-
2011	10	1,314 ± 204	1	1	-
2012	12	1,467 ± 219	1	-	-

DISCUSSION

This is the first study showing that introduction of the EWTD, which progressively reduced the average working week from 58 hours to 48 hours, has not adversely affected the operative experience of surgical residents. Results of this study are based on a retrospective analysis of operating records, involving 235,357 surgical procedures performed by a large cohort of surgeons within a surgical training region in the Netherlands, including 1 university hospital and 5 district hospitals, between 2005 and 2012. The operating records were derived from electronic hospitals systems in which all surgical procedures carried out at the operating theater complex are registered. Regular inspection of these data by external audits to evaluate the offered training opportunities ensure that these operating records are highly accurate and reliable.

After correcting for institutional volume changes, allowing a true comparison of the operative experience over the period studied, the proportional change in the number of surgical procedures performed by different grades of surgeons was examined. Interestingly, the percentage of operations carried out by attending surgeons in the university hospital declined significantly, whereas a significant increase was observed in the percentage of operations performed by fellow surgeons. Although an increase in the number of fellow surgeons most likely might have accounted for this decline, there is no data available to support this hypothesis. In addition, it must be noted that in general managing, research and educational tasks have increased significantly in recent years within our university hospital, providing an alternative explanation for the reduced percentage of operations performed by the attending surgeons.

The percentage of operations carried out by surgical residents in the university hospital did not change significantly over time (Fig. 1a). However, when examining the surgical procedures performed by residents in the district hospitals, significant differences were found. Noteworthy was the year 2008, in which the percentage of operations by surgical residents started to decrease, which was compensated by concomitant increase in the percentage of operations performed by the attending surgeons, suggesting an adverse effect of the EWTD (Fig. 1b). To assess this assumption, a more detailed analysis of the surgical procedures performed the residents was carried out, comprising an examination of the mean number of operations performed 1.0 FTE. Hence, an adjustment for potential changes in workload resulting from part-time work and maternity leave was ensured. It was found that the mean number of operations carried out by residents in the university hospital increased significantly from

128 surgical procedures in 2005 to 204 surgical procedures in 2012. This increase was associated with a significant decline in the number of FTEs employed as surgical resident in the university hospital from 21.5 in 2005 to 12.1 in 2012 (Fig. 2a), which may best be explained by the fact that the training of residents from other surgical specialties than general surgery gradually has been transferred to the district hospitals over the period studied, and currently is even no longer offered in the university hospital. For residents in the district hospitals, the mean number of operations performed per 1.0 FTE did not change significantly, and also the mean number of FTEs employed as surgical residents was stable over the period studied (Fig. 2b). In addition, the operative caseload of residents who completed the surgical training program under the work hour restrictions was examined. The results showed no significant difference over time, despite an increasing number of residents who obtained their board certification between 2005 and 2012, indicating that the operative experience of surgical residents in the Netherlands has been preserved (Table 1).

When interpreting these results, it should be considered that, along with the work hour restrictions, other developments have taken place in the Netherlands during the period studied, such as an increased influx of surgical residents at a national level, a reduction of the length of training, a shift of training posts for residents of other surgical specialties than general surgery toward the district hospitals, and implementation of a competency-based training program, which might have influenced the outcomes of this study and can be considered as potential confounding factors. Furthermore, adherence to the work hour regulations could not be assessed in this retrospective study. Therefore, non-compliance could have biased our results, especially as it is widely cited that work hour legislation is disregarded on a large scale in Europe.[18,19] However, high fines for several hospitals in the Netherlands that have violated the EWTD make this assumption less likely. In addition, it should be acknowledged that, due to the length of surgical training, the full implications of the EWTD are likely to become more obvious when a cohort of surgical residents is produced, who have exclusively been trained within a 48-hour working week. Consequently, the debate on the effect of work hour restrictions cannot be closed yet and future studies will remain necessary.

In an attempt to relate the results of this study to those of other studies, a systematic review was found in which the effect of work hour restrictions on objective measures of postgraduate medical training in both Europe and the US were examined.[25] In the majority of studies originating from the US, no change in the operative caseload of surgical residents being associated with the introduction of 80-hour working week was reported.[25] However, in Europe, studies investigating the effect of the EWTD on the operative experience of residents in surgical specialties have only been conducted in single center studies of low methodological

quality, as only residents with a certain level of training were involved, or results were not statistically analyzed.[29-34] Nonetheless, in most of those studies, implementation of the EWTD was associated with a decrease in the number of surgical procedures performed by surgical residents.

There are several explanations for the findings in this study. First, training efficiency has improved over time. Innovative surgical technologies, such as laparoscopic surgery[35] and endovascular techniques,[36] as well as the implementation of Enhanced Recovery After Surgery (ERAS) pathways,[37] have significantly reduced the length of hospital stay for surgical patients within the past decade. These developments, in combination with a more efficient use of operating theater time and an increase in day-case surgery, allow for a higher patient turnover,[38] likely leading to an increase in the number of surgical procedures performed by residents. Secondly, residents are allowed to carry out more complex surgical procedures at an earlier stage of their training due to incorporation of subspecialty training within the 6-year training program. This is ensured, without comprising patient safety, through the introduction of skills training courses and the increasing use of simulation and virtual reality devices in surgical training programs.[39] Finally, involvement of other healthcare workers, such as physician assistants and nurse practitioners, have reduced the amount of routine work formerly carried out by surgical residents.[40] As a consequence, residents can spend more time in the operating room and thus maximize the available training time in a most efficient manner, while continuity of care on the wards and outpatient clinic is covered.

It should be noted that the preservation of operative experience may come at the expense of other important aspects of training, such as attendance during clinical handovers, multidisciplinary meetings, or teaching sessions. In addition, it might be that surgical residents are less exposed to critical phases in the management of patients, such as during time of admission, decision making in the intensive care unit, grand rounds on the inpatient wards, or follow-up at the outpatient clinic.[41] In this context, Schijven et al.[42] compared practice-ready surgeons from Canada and the Netherlands, with the main difference being the number of working hours per week. They found equivalent scores for technical skills and cognitive knowledge, but in the examination for patient management skills the Canadian surgeons performed significantly better. This could imply that with the implementation of the EWTD there is an increased focus on the preservation of surgical skills, while attention paid to other important areas of training, such as development of non-technical skills and patient safety, may be compromised. Adverse events in healthcare, however, are more likely to originate from inadequate professional competencies than a lack of knowledge or technical skills.[43,44] Therefore, the future challenge is to ensure that the next generation of surgeons will continue to be

well-trained in all aspects of surgical patient care, from first meeting at the outpatient clinic or the emergency room to discharge and follow-up after surgery, within a 48-hour working week. Prolonging the duration of surgical training and a repeal of the EWTD are some of the initiatives suggested by critics to compensate for the decreased working hours.[24] The results of this study show that more efficient use of available training time may help to comply with work hour restrictions. More studies are needed to establish the effect of reduced working hours on the development of other competencies that are essential for a surgeon, in addition to surgical skills, to evaluate whether surgical residents are adequately prepared for the full scope of tasks they are expected to perform in professional practice.

REFERENCES

1. Directive 2003/88/EC of the European Parliament and of the Council of 4 November 2003 concerning certain aspects of the organisation of working time. http://eur-lex.europa.eu/smartapi/cgi/sga_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=en&numdoc=32003L0088&model=guichett [14th of January 2014].
2. Fitzgerald JE, Caesar BC. The European Working Time Directive: a practical review for surgical trainees. Int J Surg 2012;10(8): 399-403.
3. Iglehart JK. Revisiting duty-hour limits--IOM recommendations for patient safety and resident education. N Engl J Med 2008;359(25): 2633-2635.
4. Curet MJ. Resident work hour restrictions: where are we now? J Am Coll Surg 2008;207(5): 767-776.
5. Gopal R, Glasheen JJ, Miyoshi TJ, Prochazka AV. Burnout and internal medicine resident work-hour restrictions. Arch Intern Med 2005;165(22): 2595-2600.
6. Barger LK, Cade BE, Ayas NT, Cronin JW, Rosner B, Speizer FE, et al. Extended work shifts and the risk of motor vehicle crashes among interns. N Engl J Med 2005;352(2): 125-134.
7. Lockley SW, Cronin JW, Evans EE, Cade BE, Lee CJ, Landrigan CP, et al. Effect of reducing interns' weekly work hours on sleep and attentional failures. N Engl J Med 2004;351(18): 1829-1837.
8. Fletcher KE, Reed DA, Arora VM. Patient safety, resident education and resident well-being following implementation of the 2003 ACGME duty hour rules. J Gen Intern Med 2011;26(8): 907-919.
9. Jamal MH, Rousseau MC, Hanna WC, Doi SA, Meterissian S, Snell L. Effect of the ACGME duty hours restrictions on surgical residents and faculty: a systematic review. Acad Med 2011;86(1): 34-42.
10. Antiel RM, Reed DA, Van Arendonk KJ, Wightman SC, Hall DE, Porterfield JR, et al. Effects of duty hour restrictions on core competencies, education, quality of life, and burnout among general surgery interns. JAMA Surg 2013;148(5): 448-455.
11. Drolet BC, Sangisetty S, Tracy TF, Cioffi WG. Surgical residents' perceptions of 2011 Accreditation Council for Graduate Medical Education duty hour regulations. JAMA Surg 2013;148(5): 427-433.
12. Optimising Working Hours to Provide Quality in Training and Patient Safety; A Position Statement by The Association of Surgeons in Training. January 2009. http://www.asit.org/assets/documents/ASiT_EWTD_Position_Statement.pdf [14th of January 2014].
13. Morris-Stiff GJ, Sarasin S, Edwards P, Lewis WG, Lewis MH. The European Working Time Directive: One for all and all for one? Surgery 2005;137(3): 293-297.
14. Barger LK, Ayas NT, Cade BE, Cronin JW, Rosner B, Speizer FE, et al. Impact of extended-duration shifts on medical errors, adverse events, and attentional failures. PLoS Med 2006;3(12): e487.
15. Landrigan CP, Rothschild JM, Cronin JW, Kaushal R, Burdick E, Katz JT, et al. Effect of reducing interns' work hours on serious medical errors in intensive care units. N Engl J Med 2004;351(18): 1838-1848.

16. Cairns H, Hendry B, Leather A, Moxham J. Outcomes of the European Working Time Directive. BMJ 2008;337: a942.
17. Axelrod L, Shah DJ, Jena AB. The European Working Time Directive: an uncontrolled experiment in medical care and education. JAMA 2013;309(5): 447-448.
18. Doctors' training and the European Working Time Directive. Lancet 2010;375(9732): 2121.
19. Roy P. Working time regulations for trainee doctors. BMJ 2009;339.
20. Rosenbaum L, Lamas D. Residents' duty hours--toward an empirical narrative. N Engl J Med 2012;367(21): 2044-2049.
21. Glomsaker TB, Soreide K. Surgical training and working time restriction. Br J Surg 2009;96(4): 329-330.
22. Purcell Jackson G, Tarpley JL. How long does it take to train a surgeon? BMJ 2009;339: b4260.
23. Canter R. Impact of reduced working time on surgical training in the United Kingdom and Ireland. Surgeon 2011;9 Suppl 1: S6-7.
24. Benes V. The European Working Time Directive and the effects on training of surgical specialists (doctors in training): a position paper of the surgical disciplines of the countries of the EU. Acta Neurochir (Wien) 2006;148(11): 1227-1233.
25. Moonesinghe SR, Lowery J, Shahi N, Millen A, Beard JD. Impact of reduction in working hours for doctors in training on postgraduate medical education and patients' outcomes: systematic review. BMJ 2011;342: d1580.
26. Borel-Rinkes IH, Gouma DJ, Hamming JF. Surgical training in the Netherlands. World J Surg 2008;32(10): 2172-2177.
27. Hopmans CJ, den Hoed PT, Wallenburg I, van der Laan L, van der Harst E, van der Elst M, et al. Surgeons' attitude toward a competency-based training and assessment program: results of a multicenter survey. J Surg Educ 2013;70(5): 647-654.
28. Wijnhoven BP, Watson DI, van den Ende ED. Current status and future perspective of general surgical trainees in the Netherlands. World J Surg 2008;32(1): 119-124.
29. Marron CD, Byrnes CK, Kirk SJ. An EWTD-compliant shift rota decreases training opportunities. Bulletin of The Royal College of Surgeons of England 2005;87(7): 246-248.
30. Breen KJ, Hogan AM, Mealy K. The detrimental impact of the implementation of the European working time directive (EWTD) on surgical senior house officer (SHO) operative experience. Ir J Med Sci 2013;182(3): 383-387.
31. Stephens MR, Pellard S, Boyce J, Blackshaw G, Williams DH, Lewis WG. Influence of EWTD compliant rotas on SHO operative experience. Bulletin of The Royal College of Surgeons of England 2004;86(4): 120-121.
32. Kara N, Patil PV, Shimi SM. Changes in working patterns hit emergency general surgical training. Bulletin of The Royal College of Surgeons of England 2008;90(2): 60-63.

33. Maxwell AJ, Crocker M, Jones TL, Bhagawati D, Papadopoulos MC, Bell BA. Implementation of the European Working Time Directive in neurosurgery reduces continuity of care and training opportunities. Acta Neurochir (Wien) 2010;152(7): 1207-1210.
34. Oakley JE, Clewer GJ, Cool WP, Ford DJ. The Effect of the EWTD on orthopaedic specialist training. Bulletin of The Royal College of Surgeons of England 2005;87(9): 320-321.
35. Lacy AM, Garcia-Valdecasas JC, Delgado S, Castells A, Taura P, Pique JM, et al. Laparoscopy-assisted colectomy versus open colectomy for treatment of non-metastatic colon cancer: a randomised trial. Lancet 2002;359(9325): 2224-2229.
36. Bavaria JE, Appoo JJ, Makaroun MS, Verter J, Yu ZF, Mitchell RS, et al. Endovascular stent grafting versus open surgical repair of descending thoracic aortic aneurysms in low-risk patients: a multicenter comparative trial. J Thorac Cardiovasc Surg 2007;133(2): 369-377.
37. Varadhan KK, Neal KR, Dejong CH, Fearon KC, Ljungqvist O, Lobo DN. The enhanced recovery after surgery (ERAS) pathway for patients undergoing major elective open colorectal surgery: a meta-analysis of randomized controlled trials. Clin Nutr 2010;29(4): 434-440.
38. Dexter F, Traub RD. How to schedule elective surgical cases into specific operating rooms to maximize the efficiency of use of operating room time. Anesth Analg 2002;94(4): 933-942, table of contents.
39. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg 2002;236(4): 458-463; discussion 463-454.
40. Dowling S, Barrett S, West R. With nurse practitioners, who needs house officers? BMJ 1995;311(7000): 309-313.
41. Antiel RM, Thompson SM, Reed DA, James KM, Tilburt JC, Bannon MP, et al. ACGME duty-hour recommendations - a national survey of residency program directors. N Engl J Med 2010;363(8): e12.
42. Schijven MP, Reznick RK, ten Cate OT, Grantcharov TP, Regehr G, Satterthwaite L, et al. Transatlantic comparison of the competence of surgeons at the start of their professional career. Br J Surg 2010;97(3): 443-449.
43. Yule S, Paterson-Brown S. Surgeons' non-technical skills. Surg Clin North Am 2012;92(1): 37-50.
44. Kripalani S, LeFevre F, Phillips CO, Williams MV, Basaviah P, Baker DW. Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care. JAMA 2007;297(8): 831-841.

Chapter 6

Exploring the changing landscape of surgical residency training: Results of an interim multicenter survey analysis

Cornelis J. Hopmans
Pieter T. den Hoed
Reinier Timman
Jan N.M. IJzermans

On behalf of the regional training committee on surgical training representing region IV in the Netherlands

Submitted

ABSTRACT

Background: Over the past decade, surgical training has changed by the introduction of competency-based curricula, implementation of work hour restrictions, and the fragmentation of general surgery into subspecialties. The aim of this successive multicenter survey was to examine how surgical residents and attending surgeons in the Netherlands perceive these developments.

Methods: In 2011 and 2013, a survey was distributed in a surgical training region in the Netherlands, consisting of 1 university hospital and 6 district hospitals. The survey addressed the following topics: (1) introduction of a competency-based training and assessment program, (2) impact of work hour restrictions on surgical training and healthcare delivery, (3) subspecialty training and the future of general surgery. Agreement with survey items was defined as a score ≥4 on a 5-point Likert scale (indicating a positive attitude) by at least 80% of the respondents. The Mann-Whitney test was used to evaluate differences between groups.

Results: In 2011, the response rate was 140/159 (88%), at follow-up in 2013, 127/154 (82%). As before, the traditional master-apprenticeship training model is preferred over the competency-based curriculum whereas the importance of the CanMEDS competencies is recognized. The adopted assessment tools are predominantly considered unsuitable to evaluate competence progression. Attending surgeons are significantly more concerned about the consequences of work hour restrictions on surgical training and healthcare delivery as compared to surgical residents. On completion of the training program, there is a feeling of unpreparedness for independent practice as surgical specialist in any of the subspecialties offered. The future of general surgery appears to be ill-defined.

Conclusions: This interim analysis demonstrates that 4 years after the landscape of surgical residency training has been radically changed, surgical residents and attending surgeons evoke a desire to preserve traditional values and related practices, underlining the relevance of proper feedback to adjust change.

INTRODUCTION

Historically, surgical residency training has been based upon the master-apprenticeship model. In this well-established training model, characterized by the educational principles 'learning-by-doing' and 'see one, do one, teach one', the key element of success depended on a heavy clinical workload and working long hours, resulting in an enormous exposure to a broad spectrum of surgical conditions and wide repertoire of surgical procedures.[1,2] However, within the past decade, this traditional educational concept has been replaced by competency-based curricula in many Western countries.[3-6] A greater public demand for accountability and more transparency in medicine, increased awareness for patient safety due to media attention related to medical error, and calls from society to align postgraduate medical training to the altered expectations of patients are the prime drivers for this shift.7 In addition, stringent work hour restrictions have been progressively enforced by law, substantially reducing surgical residents' training time and availability during daytime. Within the medical profession, and in particular surgical specialties, these work hour restrictions are believed to have deteriorated the quality of training and adversely affect patient safety.[8-10] Another important development having a profound impact on surgical residency training is the progressive fragmentation of general surgery into subspecialties. Innovative surgical techniques in combination with an increasingly well-informed public that prefers advanced specialized care, as well as evidence demonstrating that centralization of particular surgical procedures leads to improved patient outcomes are all factors enhancing the upward trend toward subspecialty training.[11,12]

Over the past years, also in the Netherlands, the structure and format of surgical residency training has changed radically. In 2009, as part of a national initiative to reform postgraduate medical training, a competency-based curriculum based on 7 competencies of the Canadian Medical Education Directives for Specialists (CanMEDS) framework (Medical expert, Communicator, Collaborator, Scholar, Health Advocate, Manager and Professional) was introduced.[13] To document and evaluate competence progression in an objective and standardized way, various assessment tools have been adopted on which scores and feedback are assigned, referring to the different CanMEDS competencies.[14-18] As a result, assessment tools, which are collected in a portfolio, provide a transparent overview of surgical residents' progression in achieving predefined competency goals. In Table 1, an overview of the adopted assessment tools can be seen, showing which competencies can be evaluated using the different assessment tools. In addition, implementation of the European Working Time Directive (EWTD) has progressively reduced surgical residents' working hours, limiting the average working week

Table 1. Description of the assessment tools introduced by the Association of Surgeons of the Netherlands to evaluate the CanMEDS competencies.[45]

Assessment Tools	Definition	CanMEDS competencies						
		ME	COM	COL	SCH	HA	MAN	PRO
In-training evaluation report	Periodic review of the portfolio for completeness and evaluation of competence progression	X	X	X	X	X	X	X
360-degree assessment	Evaluation of daily functioning by colleagues, nursing staff and other healthcare workers		X	X				X
Critically appraised topic	Critical appraisal of a scientific article relating to a focused clinical question derived from a real-patient encounter				X			
Case-based discussion	Evaluation of clinical reasoning and decision-making skills after the presentation of a managed clinical case	X			X			
mini-Clinical evaluation exercise	Observation and assessment of clinical skills, attitudes and behaviour during patient encounters, clinical handovers or multidisciplinary meetings	X	X	X		X		X
Assessment of an operation report	Critical assessment of an operation report	X					X	X
Report of a complicated hospitalization	Review of a patient record to evaluate medical record keeping and correspondence	X					X	X
Objective structured assessment of technical skills	Assessment of surgical residents' operative skills in the operating theater.	X						

ME (Medical Expert); COM (Communicator); COL (Collaborator); SCH (Scholar); HA (Health Advocate); MAN (Manager); PRO (Professional)

to a maximum of 48 hours.[19] On top of this significant reduction in training time, training in general surgery, often lasting 8 years due to postresidency fellowship training, has been shortened to a 6-year program in which subspecialty training is mandatory and offered during the last 2 years. As a consequence, surgical residents are currently no longer trained as broad-based general surgeons, but as surgical specialists who will primarily be focused on the subspecialty in which training has been received for the rest of their professional careers, such as gastro-intestinal surgery, surgical oncology, trauma surgery or vascular surgery.[20]

An important aspect of implementing a new training methodology includes monitoring and evaluation. Data obtained from such evaluations provide important feedback about user satisfaction, identify the need for adjustments, and may be used to improve surgical residency training and education. Furthermore, information regarding other important developments, such as the implementation of work hour restrictions and the fragmentation of general surgery into subspecialties, is urgently needed, potentially providing valuable insights for the future workforce planning and organization of surgical training and healthcare delivery. The aim of this successive multicenter survey was to provide an interim analysis on how surgical residents and attending surgeons in the Netherlands perceive the transformation of surgical training, awaiting the next repeated measurement scheduled to be conducted later this year.

METHODS

Setting and Study Design

In the Netherlands, training in general surgery is regionally organized. Each training region (n=8) consists of a university hospital and several affiliated district hospitals. This study was conducted in a region in which surgical training is organized in 1 university hospital and 6 district hospitals. In November 2011, all surgical residents and attending surgeons were invited to complete a structured survey, which was distributed in a paper-based format together with a reply-paid envelope, as well as electronically using an e-mail with a web-based link (www.surveymonkey.com). An introductory letter enclosed with the survey explained the purpose of the study, informed that participation was voluntary and assured that data would be treated confidentially and anonymously. If surveys had not been returned within 2 weeks, non-responders were sent a reminder by e-mail. The maximum response time was 1 month. The same methodology was used for surveys distributed at follow-up in November 2013. Ethical approval for this study was not required.

Survey

The survey was based on semi-structured interviews that were conducted in 2011 among surgical residents (n=12) and attending surgeons (n=14). More detailed information about the development of the survey can be found in a recently published article by our research team.[21] The survey addressed the following topics: (1) introduction of a competency-based curriculum, importance of the various CanMEDS roles for the competency profile of a surgeon and suitability of the adopted assessment tools to evaluate competence development, (2) impact of work hour restrictions on surgical training and healthcare delivery, and (3) subspecialty training and the future of general surgery. Questions had to be scored on a 5-point Likert scale; score 1= strongly disagree, 2= disagree, 3= neutral, 4= agree, 5 = strongly agree. Finally, respondents were asked to indicate gender, age, seniority (position as attending surgeon or surgical resident) and type of hospital employment.

Statistical Analysis

Data were analyzed using IBM SPSS version 21.0 (IBM corp., Armonk, NY, US). Descriptive statistics were calculated for all questions, including means, standard deviations (SD), frequencies, and ranges. In accordance with expert guidelines, agreement with survey items was defined as a score ≥4 on the 5-point Likert scale (indicating a positive attitude) by at least 80% of the respondents.[22,23] The Mann-Whitney test was used to compare responses from the group of surgical residents with those of the group of attending surgeons, as well as to evaluate whether differences in attitude within these groups occurred over the period studied. The Bonferroni adjustment was applied to correct for multiple testing, specifying the level for statistical significance at $p<0.001$.

RESULTS

In 2011, the overall response rate was 140/159 (88%), consisting of 43/51 (84%) surgical residents and 97/108 (90%) attending surgeons. At follow-up in 2013, 127 of the 154 (82%) invited participants completed the survey, including 44/51 (86%) surgical residents and 83/103 (81%) attending surgeons. Demographic characteristics of both these groups can be seen in Table 2.

Table 2. Demographic data of the respondents.

	Surgical Residents		Attending Surgeons	
	2011	**2013**	**2011**	**2013**
Number of respondents	43 (84)	44 (86)	97 (90)	83 (81)
Female	17 (40)	19 (43)	13 (13)	12 (15)
Male	26 (60)	25 (57)	84 (87)	71 (86)
Age in years, mean (range):	31.6 (27-36)	31.4 (28-36)	44.9 (33-63)	47.1 (33-65)
University hospital	9 (21)	10 (23)	33 (34)	23 (28)
District hospital	34(79)	34 (77)	64 (66)	60 (72)

Data are presented as number (percentage) unless otherwise indicated

Competency-based curriculum

In 2011, two years after the educational concept of the surgical training program was changed, most surgical residents (77%) and attending surgeons (83%) indicated to be aware of the content of the competency-based curriculum. In 2013, this percentage increased to 86% for the group of residents, whereas a similar percentage was found for the group of attending surgeons. Initially, 44% of the surgical residents and 29% of the attending surgeons considered the competency-based curriculum an appropriate training model, with percentages increasing to 51% and 40% in 2013, respectively. In both 2011 and 2013, a clear majority of the surgical residents and the attending surgeons considered the master-apprentice relationship as a key element in surgical training. In 2011, only a small minority of the surgical residents (14%) and the attending surgeons (22%) indicated that competency-based training is a better educational principle than the traditional master-apprenticeship model. In 2013, these percentages increased to 41% and 28%, respectively (Table 3a). In 2011, the vast majority of the surgical residents and the attending surgeons recognized the importance of most CanMEDS competencies, with exception of the competencies Manager (surgical residents 76%; attending surgeons 78%) and Health Advocate (surgical residents 69%; attending surgeons 70%. In 2013, the importance of the CanMEDS competencies was confirmed, with only the competency Health Advocate failing to achieve agreement percentages ≥80% (surgical residents 60%; attending surgeons 61%). These results can be seen in Table 3b. With agree-

Table 3a-b-c. Summary of data on respondents' attitude toward the introduction of a competency-based curriculum: (A) competency-based training program (B) CanMEDS competencies
(C) assessment tools.

	2011						2013					
	Surgical Residents			Attending Surgeons			Surgical Residents			Attending Surgeons		
A) Competency-based training program	**% (4/5)**	**Mean**	**SD**	**% (4/5)**	**Mean**	**SD**	**% (4/5)**	**Mean**	**SD**	**% (4/5)**	**Mean**	**SD**
I am informed about the content of the competency-based curriculum	77	3.91	0.75	83	3.90	0.76	86	3.98	0.63	83	3.98	0.72
The competency-based curriculum is an appropriate training model	44	3.35	0.72	29	3.21	0.64	51	3.47	0.67	40	3.26	0.77
The master-apprentice relationship is a key element in surgical training	86	4.26	0.69	78	3.89	0.87	84	4.09	0.71	86	4.19	0.74
Competency-based training is a better educational concept than the traditional master-apprenticeship model	14	2.93	0.67	22	2.93	0.84	41	3.27	0.85	28	2.92	0.94
B) CanMEDS competencies	**% (4/5)**	**Mean**	**SD**	**% (4/5)**	**Mean**	**SD**	**% (4/5)**	**Mean**	**SD**	**% (4/5)**	**Mean**	**SD**
Medical Expert	98	4.74	0.45	99	4.78	0.55	100	4.71	0.46	100	4.65	0.48
Communicator	98	4.67	0.48	99	4.53	0.61	100	4.56	0.50	99	4.50	0.53
Collaborator	98	4.55	0.50	99	4.51	0.61	100	4.57	0.51	99	4.43	0.53
Scholar	93	4.29	0.60	95	4.27	0.64	90	4.18	0.59	93	4.27	0.58
Health Advocate	69	3.83	0.73	70	3.74	0.68	60	3.60	0.67	61	3.70	0.81
Manager	76	4.02	0.72	78	3.93	0.71	83	4.00	0.60	81	3.97	0.65
Professional	98	4.45	0.55	98	4.50	0.63	100	4.40	0.59	96	4.43	0.62
C) Assessment tools	**% (4/5)**	**Mean**	**SD**	**% (4/5)**	**Mean**	**SD**	**% (4/5)**	**Mean**	**SD**	**% (4/5)**	**Mean**	**SD**
In-Training Evaluation Report (ITER)	95	4.28	0.55	90	4.07	0.57	93	4.24	0.66	94	4.22	0.55
360-degree Assessment	54	3.56	0.91	64	3.70	0.85	62	3.62	0.99	68	3.71	0.77
Critically Appraised Topic (CAT)	47	3.30	0.74	59	3.59	0.64	45	3.29	0.84	59	3.51	0.65
Case-based Discussion (Cbd)	49	3.32	0.79	70	3.70	0.65	46	3.30	0.90	66	3.69	0.57
Mini-Clinical Evaluation exercise (Mini-CEX)	54	3.42	0.88	70	3.66	0.75	59	3.51	0.95	75	3.73	0.82
Assessment of an Operation Report	54	3.42	0.93	56	3.45	0.78	59	3.46	0.87	60	3.38	0.86
Report of a Complicated Hospitalization	47	3.23	0.90	58	3.63	0.76	49	3.32	0.83	77	3.80	0.60
Objective Structured Assessment of Technical Skills (OSATS)	84	4.09	0.72	81	3.99	0.73	93	4.19	0.55	82	3.94	0.75
Assessment tools cause an administrative burden	81	4.02	0.83	87	4.20	0.72	68	3.79	0.77	71	3.89	0.85
Assessment tool cannot replace the opinion supervising surgeons have regarding the functioning of residents based on close cooperation	72	3.98	0.91	71	3.83	0.91	81	4.09	0.87	73	3.87	0.87
Assessment tools are helpful to structure and coordinate feedback	79	3.81	0.76	83	3.94	0.60	58	3.47	0.86	82	3.91	0.63
The keeping of a portfolio is meaningful.	77	3.88	0.59	74	3.76	0.59	77	3.81	0.59	76	3.84	0.63

% 4/5 -> Proportion of respondents rating a survey item with a score of 4 or 5 (on a 5-point Likert scale). Data are presented as means and standard deviations (SD). Over the period studied, no significant differences were found in the responses between the surgical residents and the attending surgeons. Within both the group of respondents, also no significant change in attitude was observed (data not shown).

ment percentages ≥80% in both 2011 and 2013, only the instruments In-Training Evaluation Report (ITER) and Objective Structured Assessment of Technical Skills (OSATS) of the various assessment tools adopted were perceived suitable for competence evaluation purposes. In 2011, most of the surgical residents (81%) and attending surgeons (87%) blamed the assessment tools for causing an administrative burden. In 2013, these percentages decreased to 68% and 71%, respectively. Furthermore, it was felt that assessment tools cannot replace the opinion attending surgeons have regarding the functioning of surgical residents based on close cooperation, with percentages of 71% and 72% for the different groups of respondents in 2011, slightly increasing to 73% and 81% in 2013. In both 2011 and 2013, most of the surgical residents and attending surgeons indicated that assessment tools are helpful to structure and coordinate feedback. The keeping of a portfolio was considered useful by the group of surgical residents (74%) and the group of attending surgeons (77%) in 2011. In 2013 almost similar percentages were found for both groups. These results can be seen in Table 3c. With respect to the questions related to the introduction of a competency-based training and assessment program, no significant differences were found in the responses between the surgical residents and attending surgeons in both 2011 and 2013. Also within the group of surgical residents and the group of attending surgeons, no significant change in attitude could be measured over the period studied (data not shown).

Work hour restrictions

In 2011, 47% of the surgical residents and 26% of the attending surgeons agreed that the 48-hour maximum working week provides sufficient time for adequate training, followed by an increase to 58% and 41% in 2013, respectively. In 2011, 54% of the surgical residents and 61% of the attending surgeons indicated that the 48-hour working has improved work-life balance. Almost similar scores were found for both groups in 2013. Most of the surgical residents (68%) and the attending surgeons (78%) indicated that work hour restrictions reduce the operative caseload of surgical residents, with percentages decreasing to 43% and 60% in 2013. Other questions exploring respondents' attitude toward the impact of work hour restrictions revealed that attending surgeons were more concerned about the consequences of work hour restrictions compared to the surgical residents in both 2011 and 2013 with respect to the continuity of surgical training ($p<0.001$), the number of patient encounters during training ($p<0.001$), and the continuity of healthcare delivery ($p<0.001$). These results can be seen in Table 4. Comparing responses from the group of surgical residents over the period studied revealed no significant differences. However, in 2013, the group of attending surgeons assigned lower scores to questions examining the perceived consequences of the work hour restrictions on surgical residents' operative caseload (2011 mean=3.85 versus 2013 mean=3.41; $p<0.001$) and the continuity of healthcare delivery (2011 mean=4.16 versus 2013 mean=3.53; $p<0.001$).

Table 4. Summary of data on respondents' attitude regarding the impact of work hour restrictions on surgical residency training and healthcare delivery.

	2011							2013						
	Surgical Residents			Attending Surgeons				Surgical Residents			Attending Surgeons			
Work hour restrictions	**% (4/5)**	**Mean**	**SD**	**% (4/5)**	**Mean**	**SD**	**p-value**	**% (4/5)**	**Mean**	**SD**	**% (4/5)**	**Mean**	**SD**	**p-value**
An 48-hour working week provides sufficient time for adequate training	47	3.09	1.02	26	2.76	0.94	0.051	58	3.37	0.95	41	3.15	0.93	0.163
Work hour restrictions have improved surgical residents' work-life balance	54	3.40	1.22	61	3.59	0.78	0.557	58	3.57	0.83	58	3.58	0.77	0.983
Work hour restrictions reduce the operative caseload of surgical residents	68	3.65	0.92	78	3.85	0.88	0.199	43	3.14	1.05	60	3.41	0.90	0.155
The continuity of surgical training is deteriorated by the work hour restrictions	49	3.35	1.20	84	3.99	0.77	<0.001	26	2.88	0.94	67	3.74	0.81	<0.001
Work hour restrictions decrease the number of patient encounters by surgical residents	49	3.30	1.08	83	3.95	0.80	<0.001	27	2.79	0.98	68	3.67	0.85	<0.001
Work hour restrictions adversely affect the continuity of healthcare delivery	56	3.49	1.08	92	4.16	0.59	<0.001	31	2.86	1.03	60	3.53	1.02	<0.001

% 4/5 -> Proportion of respondents rating a survey item with a score of 4 or 5 (on a 5-point Likert scale). Data are presented as means and standard deviations (SD).

Subspecialty training and future of general surgery

In both 2011 and 2013, most of the surgical residents and attending surgeons agreed that subspecialty training improves the quality of patient care. Whether it will be feasible to produce competent surgical specialists after 6 years of training remains to be seen, with respondents' percentages for the different subspecialties offered ranging from 37% to 56% in 2011, and from 19% to 51% in 2013. In both 2011 and 2013, only a small minority of the surgical residents and the attending surgeons indicated that postresidency fellowship training will be superfluous in the future. Extending the duration of subspecialty training to 4 years was not supported by the surgical residents (30%) or by the attending surgeons (35%). In 2013, almost similar percentages were found for both groups (Table 5a). As shown in Table 5b, opinions regarding the future of general surgery were mixed. In 2011, 47% of the surgical residents and 63% of the attending surgeons indicated that general surgeons will continue to be needed in the future. In 2013, almost similar responses were received from both groups. In both 2011 and 2013, conflicting feelings were also expressed regarding to the question whether general surgery should be added as a subspecialty in the curriculum. Over the period studied, no significant differences were found in the responses between the surgical residents and the attending surgeons. Within both the group of respondents, also no significant change in attitude was observed (data not shown).

Table 5a-b. Summary of data on respondents' attitude toward subspecialty training (A) and the future of general surgery (B).

	2011						2013					
	Surgical Residents			Attending Surgeons			Surgical Residents			Attending Surgeons		
A) Subspecialty training	% (4/5)	Mean	SD	% (4/5)	Mean	SD	% (4/5)	Mean	SD	% (4/5)	Mean	SD
Subspecialty training improves the quality of patient care	74	3.77	1.04	73	3.97	0.88	77	3.77	0.75	76	3.86	0.77
A competent vascular surgeon can be trained in 6 years	44	3.12	1.14	47	3.19	1.13	19	2.83	1.03	44	3.12	1.14
A competent trauma surgeon can be trained in 6 years	54	3.42	1.12	56	3.38	1.06	29	3.02	0.98	51	3.29	1.07
A competent gastrointestinal surgeon can be trained in 6 years	42	3.09	1.11	51	3.30	1.09	28	2.93	1.01	44	3.18	1.10
A competent surgical oncologist can be trained in 6 years	37	3.02	1.06	47	3.20	1.08	21	2.81	0.98	42	3.06	1.16
Postresidency fellowship training will be superfluous in the future	14	2.26	0.98	11	2.16	0.97	14	2.45	0.99	9	2.10	0.92
The duration of subspecialty training should be extended to 4 years (2+4 schedule)	30	2.60	1.24	35	2.82	1.26	40	2.93	1.16	33	2.91	1.15
B) Future of general surgery	% (4/5)	Mean	SD	% (4/5)	Mean	SD	% (4/5)	Mean	SD	% (4/5)	Mean	SD
General surgeons will continue to be needed in the future	47	3.33	1.04	63	3.59	1.14	61	3.58	0.88	58	3.56	1.04
General surgery should be included as a subspecialty in the curriculum	48	3.26	1.04	51	3.34	1.19	42	3.16	1.00	45	3.32	1.12

% 4/5 -> Proportion of respondents rating a survey item with a score of 4 or 5 (on a 5-point Likert scale). Data are presented as means and standard deviations (SD). Over the period studied, no significant differences were found in the responses between the surgical residents and the attending surgeons. Within both the group of respondents, also no significant change in attitude was observed (data not shown).

DISCUSSION

Surgical training has undergone extensive reform within the past decade. Previous studies examining the impact of educational reform have been primarily conducted using a cross-sectional study design. Yet, it is unknown whether a change in attitude has occurred over time.[24] The aim of this successive multicenter survey was to provide an interim analysis on how surgical residents and attending surgeons perceive the transformation of surgical training, allowing to track changes in opinions and attitudes over time. Each of the topics addressed in the survey has revealed several important findings. This information may be used for continuous improvement of surgical curricula and may be helpful for stakeholders to guide future directions in surgical training and healthcare delivery.

In most Western countries, the educational concept of surgical residency training has shifted from the master-apprenticeship model toward competency-based curricula. The results of this study, however, show that both surgical residents and attending surgeons still prefer the traditional apprenticeship training model over a competency-based curriculum. Over the period studied, no significant change in attitude was measured. This finding is not surprising given the well-established established training paradigm that has produced competent surgeons for many decades. Also in other countries, stakeholders responsible for the organization of surgical residency training, struggle and encounter resistance to effectively integrate modern training and teaching methods.[24,25]

An important aspect of competency-based medical education is that, in addition to medical-technical skills, development of general competencies play a pivotal role. Effectively working within a team, adequate communication and information sharing with others involved in the patient's care, thus ensuring patient safety, attention for quality improvement, as well as respect for the patient's perspective in decision making processes are some examples of these general competencies and can all be related to the CanMEDS competencies. In this study, both the surgical residents and the attending surgeons generally recognized the importance of the CanMEDS competencies, especially of those in addition to Medical Expert. This is an important finding as it has been shown that surgeons who run into problems in their professional lives rarely do so because of lack of medical knowledge or poor operative skills. In most cases, problems are caused by inadequate professional competence.[26,27] Future studies will be needed to explore how teaching strategies in general skills best can be integrated in daily clinical practice.

In competency-based curricula, assessment tools are essential to confirm in an objective, standardized and transparent way whether a certain level of competence has been achieved. In this study, the adopted assessment tools were predominantly perceived unsuitable to evaluate competence progression. The instruments ITER and OSATS were an exception. This finding is not surprising as these instruments reflect to the way surgeons were traditionally assessed, based on subjective conversations and indicators of productivity. The negative perception regarding the assessment tools might be explained the administrative burden associated with completing these evaluations. It remains to be demonstrated whether the recent switch from a paper-based portfolio to a digital format will reduce the administrative burdens. In addition, it should be noted that no high-level evidence is available that the abundance of assessment tools developed can provide an accurate evaluation of surgical residents' skills.[27-29] The process of designing and verifying the reliability and validity of an assessment tool is complex and, in part, may explain the lack of appropriate evaluation methods currently available. Most of the surgical residents and attending surgeons indicated that assessment tools are helpful to structure and coordinate feedback, underscoring that there is a need optimize the receipt of feedback in surgical residency training.

The aim of work hour restrictions is to protect the health and safety of healthcare professionals, as well as to reduce the risk to patient safety posed exhaustion and fatigue resulting from working excessive long hours. Several studies have demonstrated that surgical residents' personal well-being and lifestyle have improved since the introduction of work hour restrictions.[30,31] In contrast, the results of this study demonstrate that only a slight majority of the respondents considered that reduced working hours have improved work-life balance. A potential explanation for this finding is that reduction of the working week leads to an increase in irregular hours of shift-work and complex rotation schedules, which may be perceived as more tiring and conflicting with social activities and family life.

Within the surgical profession, work hour restrictions are in general being condemned for having a detrimental effect on the quality of training and healthcare delivery.[32] The results of this study underline these contentions. The attending surgeons expressed significantly more concern about the consequences of work hour restrictions on the continuity of training and healthcare delivery as compared to the surgical residents. Over the period studied, this concern diminished significantly, reflecting that work hour restrictions are becoming more accepted. In the past years, the impact of reduced working hours on the quality of training has been extensively examined, showing that the operative caseload of surgical residents has not been adversely affected.[33,34] However, it should be noted that case-numbers alone are not the surrogate of a well-trained surgeon. The impact of work hour restrictions on

the continuity of patient care remains poorly understood with studies reporting conflicting results and theories.[35,36] Consequently, discussions on the effect of work hour restrictions are ongoing and future studies will remain necessary.

Historically, progressive specialization occurred gradually over the course of a professional career, narrowing a surgeon's scope of practice. Today, an explicit period of subspecialty training is incorporated and a mandatory part in most surgical training programs. Factors enhancing the upward trend toward subspecialty training include the exponential expansion of medical knowledge, the rapid adoption of innovative surgical techniques, evidence demonstrating that patient outcomes are improved if performed by experts in centers of excellence, as well as an increasingly well- informed public seeking advanced and specialized healthcare.[37] The results of this study also support the growth of this phenomenon. Most of the surgical residents and attending surgeons indicated that subspecialty training improves the quality of patient care.

Over the past years, general surgery is fragmented into various subspecialties that are progressively transforming from areas of interest within general surgery into recognized monospecialties with their own training program and standards.[38,39] In the Netherlands, training in general surgery, often lasting 8 years due to postresidency fellowship training, has been shortened to a 6-year program. During the first 4 years, training in general surgery is offered. The remaining 2 years consist of dedicated training in one of the following subspecialties: gastro-intestinal surgery, surgical oncology, trauma surgery or vascular surgery. The surgical residents and attending surgeons participating in this study were asked to indicate the feasibility to produce competent surgical specialists after 6 years of training. The results quite clearly demonstrate that within both groups of respondents a feeling of unpreparedness for independent practice as surgical specialist prevails for each of the subspecialties offered. This finding is consistent with the results of a questionnaire distributed among subspecialty program directors in North America reporting a perceived lack of readiness of general surgery residents to enter independent surgical practice.[40]

Taking into account the concerns regarding the preparation of surgical residents for independent practice, it is not surprisingly that both groups of respondents rejected the statement that postresidency fellowship training will be superfluous in the future. In this context, a study exploring the job market for general surgeons in the US showed that most general surgery residents pursue fellowship training, despite the fact that the majority of available jobs do not require fellowship training.[41] This finding has important implications for the future workforce planning and organization of surgical healthcare, with the next generation

of surgeons being unlikely to encompass the same depth, breadth and availability of services that general surgeons historically provided. To ensure that patients have equal access to quality healthcare facilities, training surgeons that match today's practice needs is critical

Remarkably, extending the duration of subspecialty training to 4 years, providing more time to gain experience and develop competence in the subspecialty of choice, was not supported by the surgical residents or the attending surgeons. As the objective of surgical training is to ensure optimal preparation of surgical residents for independent professional practice, we do not share this view and suggest that remodeling of the current training schedule should be considered. After 2 years of common trunk, a period shared with other surgical specialties aimed at teaching the overarching principles of medical practice and basic operative skills, surgical residents should spend the final phase of training on subspecialty training alone. As a consequence, investment in detailed medical knowledge, as well as complex surgical skills and techniques that have no relevance for the intended practice, will be avoided. Completion of such a training pathway would produce highly-qualified surgeons within a 6-year time frame that are adequately prepared for the full scope of tasks being expected to perform in an increasing complex healthcare environment in which surgical care is practiced in a multidisciplinary setting with competing demands.

The future of general surgery is a controversial topic. While some do not feel that radical changes are necessary, many believe it is mandatory. Among those who agree that surgical training should be altered, the methods of change are not universally agreed upon.[42] Also in this study, the future of general surgery appears to be ill-defined, with both surgical residents and attending surgeons expressing conflicting feelings. In literature, the views on the future of training in general vary greatly depending on various perspectives, such as geographic considerations (rural vs. urban) or practice desires (generalist vs. specialist). Also, personal factors, such as lifestyle expectations, significantly affect career planning and stimulate the tendency of specialization.[43,44] Therefore, we propose that surgical training should be tailored to market demands in order to increase access to surgical services, improve quality, and contain healthcare costs. Obviously, it is the responsibility of the surgical profession to ensure that properly trained surgeons are produced that are committed and have competency in the skills to optimally serve all kinds of patients.

This study has several limitations that must be acknowledged. This study was conducted in a single training region of the Netherlands. In addition, a disadvantage of survey methodology is that results reflect the subjective opinions of the respondents. However, in view of the high response rate and the multicenter approach, comprising 1 university hospital and 6 district

hospitals, and the fact that qualitative data served as the basis for this survey, allowing to generate in-depth insight and obtain different perspectives on how surgical residents and attending surgeons within this particular training region perceive contemporary developments in surgical training, we have no reason to believe that the findings in this study would be incongruent with the opinions expressed in other surgical training regions in the Netherlands. Furthermore, it might be possible that this interim analysis has been conducted to soon after the transformation of surgical residency training, with the competency-based curriculum and subspecialty training both being introduced in 2009 and the full implementation of the EWTD in 2011. Since more data over longer periods of time allow for more concise results, a longitudinal follow-up of this study would reveal how the attitude of surgical residents and attending surgeons toward the transformation of surgical residency training has evolved over time.

In conclusion, this successive multicenter survey analysis explored the attitude of surgical residents and attending surgeons toward the changing landscape of surgical residency training. The results demonstrate that 4 years after surgical residency training has radically been changed, contemporary developments evoke a desire to preserve traditional values and related practices, underlining the relevance of proper feedback to adjust change. The outcomes of the 6-year follow-up have to be awaited to define more precisely whether the attitude toward the transformation of surgical residency training has evolved over the years .

ACKNOWLEDGMENTS

The authors thank the members of the training committee on surgical training representing region IV in the Netherlands, consisting of Lijckle van der Laan, Amphia Hospital, Breda; Erwin van der Harst, Maasstad Hospital, Rotterdam; Maarten van der Elst, Reinier de Graaf Groep, Delft; Imro Dawson, IJsselland Hospital, Capelle aan den IJssel; Bas P.L. Wijnhoven, Erasmus MC University Medical Center, Rotterdam

REFERENCES

1. Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. BMJ 2003;327(7422): 1032-1037.
2. Reznick RK. Teaching and testing technical skills. Am J Surg 1993;165(3): 358-361.
3. Collins JP, Civil ID, Sugrue M, Balogh Z, Chehade MJ. Surgical education and training in Australia and New Zealand. World Journal of Surgery 2008;32(10): 2138-2144.
4. Frank JR, Langer B. Collaboration, communication, management, and advocacy: teaching surgeons new skills through the CanMEDS Project. World J Surg 2003;27(8): 972-978; discussion 978.
5. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system--rationale and benefits. N Engl J Med 2012;366(11): 1051-1056.
6. Smith AJ, Aggarwal R, Warren OJ, Paraskeva P. Surgical training and certification in the United kingdom. World J Surg 2009;33(2): 174-179.
7. Jones R, Higgs R, de Angelis C, Prideaux D. Changing face of medical curricula. Lancet 2001;357(9257): 699-703.
8. Axelrod L, Shah DJ, Jena AB. The European Working Time Directive: an uncontrolled experiment in medical care and education. JAMA 2013;309(5): 447-448.
9. Glomsaker TB, Soreide K. Surgical training and working time restriction. Br J Surg 2009;96(4): 329-330.
10. Rosenbaum L, Lamas D. Residents' duty hours--toward an empirical narrative. N Engl J Med 2012;367(21): 2044-2049.
11. Borman KR, Vick LR, Biester TW, Mitchell ME. Changing demographics of residents choosing fellowships: longterm data from the American Board of Surgery. J Am Coll Surg 2008;206(5): 782-788; discussion 788-789.
12. Birkmeyer JD, Siewers AE, Finlayson EV, Stukel TA, Lucas FL, Batista I, et al. Hospital volume and surgical mortality in the United States. N Engl J Med 2002;346(15): 1128-1137.
13. Frank JR. The CanMEDS 2005 Physician Competency Framework. Better standards. Better Physicians. Better care. Ottawa: The Royal College of Physicians and Surgeons of Canada.
14. Lockyer J. Multisource feedback in the assessment of physician competencies. J Contin Educ Health Prof 2003;23(1): 4-12.
15. 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.
16. Norcini JJ, Blank LL, Arnold GK, Kimball HR. The mini-CEX (clinical evaluation exercise): a preliminary investigation. Ann Intern Med 1995;123(10): 795-799.

17. Norman G, Davis D, Painvin A, Lindsay E, Rath D, Ragbeer M. Comprehensive assessment of clinical competence of family/general physicians using multiple measures. Proceedings of the Annual Conference on Research in Medical Education Washington DC: Association of American Medical Colleges 1989: 75-79.
18. Sauve S, Lee HN, Meade MO, Lang JD, Farkouh M, Cook DJ, et al. The Critically Appraised Topic: A Practical Approach to Learning Critical Appraisal. Annals RCPSC 1995;28(7): 396-398.
19. Fitzgerald JE, Caesar BC. The European Working Time Directive: a practical review for surgical trainees. Int J Surg 2012;10(8): 399-403.
20. Borel-Rinkes IH, Gouma DJ, Hamming JF. Surgical training in the Netherlands. World J Surg 2008;32(10): 2172-2177.
21. Wallenburg I, Hopmans CJ, Buljac-Samardzic M, den Hoed PT, Ijzermans JNM. Repairing reforms and transforming professional practices: a mixed-methods analysis of surgical training reform. Journal of Professions and Organization 2016;3(1): 86-102.
22. Grant JS, Davis LL. Selection and use of content experts for instrument development. Res Nurs Health 1997;20(3): 269-274.
23. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health 2006;29(5): 489-497.
24. Mehmood S, Anwar S, Ahmed J, Tayyab M, O'Regan D. A survey of UK surgical trainees and trainers; latest reforms well understood but perceived detrimental to surgical training. Surgeon 2012;10(1): 9-15.
25. Yeo H, Viola K, Berg D, Lin Z, Nunez-Smith M, Cammann C, et al. Attitudes, training experiences, and professional expectations of US general surgery residents: a national survey. JAMA 2009;302(12): 1301-1308.
26. Yule S, Paterson-Brown S. Surgeons' non-technical skills. Surg Clin North Am 2012;92(1): 37-50.
27. Kripalani S, LeFevre F, Phillips CO, Williams MV, Basaviah P, Baker DW. Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care. JAMA 2007;297(8): 831-841.
28. Carraccio C, Wolfsthal SD, Englander R, Ferentz K, Martin C. Shifting paradigms: from Flexner to competencies. Acad Med 2002;77(5): 361-367.
29. Lurie SJ, Mooney CJ, Lyness JM. Measurement of the general competencies of the accreditation council for graduate medical education: a systematic review. Acad Med 2009;84(3): 301-309.
30. Curet MJ. Resident work hour restrictions: where are we now? J Am Coll Surg 2008;207(5): 767-776.
31. Jamal MH, Rousseau MC, Hanna WC, Doi SA, Meterissian S, Snell L. Effect of the ACGME duty hours restrictions on surgical residents and faculty: a systematic review. Acad Med 2011;86(1): 34-42.
32. Doctors' training and the European Working Time Directive. Lancet 2010;375(9732): 2121.
33. Hopmans CJ, den Hoed PT, van der Laan L, van der Harst E, van der Elst M, Mannaerts GH, et al. Impact of the European Working Time Directive (EWTD) on the operative experience of surgery residents. Surgery 2015;157(4): 634-641.

34. Moonesinghe SR, Lowery J, Shahi N, Millen A, Beard JD. Impact of reduction in working hours for doctors in training on postgraduate medical education and patients' outcomes: systematic review. BMJ 2011;342: d1580.
35. Barger LK, Ayas NT, Cade BE, Cronin JW, Rosner B, Speizer FE, et al. Impact of extended-duration shifts on medical errors, adverse events, and attentional failures. PLoS Med 2006;3(12): e487.
36. Cairns H, Hendry B, Leather A, Moxham J. Outcomes of the European Working Time Directive. BMJ 2008;337: a942.
37. Stitzenberg KB, Sheldon GF. Progressive specialization within general surgery: adding to the complexity of workforce planning. J Am Coll Surg 2005;201(6): 925-932.
38. Bell PR. Vascular surgery: recognition at last. Vascular 2012;20(5): 241-242.
39. Soreide K. Trauma and the acute care surgery model--should it embrace or replace general surgery? Scand J Trauma Resusc Emerg Med 2009;17: 4.
40. Mattar SG, Alseidi AA, Jones DB, Jeyarajah DR, Swanstrom LL, Aye RW, et al. General surgery residency inadequately prepares trainees for fellowship: results of a survey of fellowship program directors. Ann Surg 2013;258(3): 440-449.
41. Decker MR, Bronson NW, Greenberg CC, Dolan JP, Kent KC, Hunter JG. The general surgery job market: analysis of current demand for general surgeons and their specialized skills. J Am Coll Surg 2013;217(6): 1133-1139.
42. Richardson JD. Training of general surgical residents: what model is appropriate? Am J Surg 2006;191(3): 296-300.
43. Fischer JE. The impending disappearance of the general surgeon. JAMA 2007;298(18): 2191-2193.
44. Debas HT. Surgery: a noble profession in a changing world. Ann Surg 2002;236(3): 263-269.
45. Structuur Curriculum Heelkunde voor Reflectieve Professionals (SCHERP). General Surgical Residency Training Program in the Netherlands www.heelkunde.nl [accessed 31-9-2013.

Chapter 7

Identification of factors associated with performance during surgical training

Cornelis J. Hopmans
Susanna M. Lucieer
Pieter T. den Hoed
Reinier Timman
Rosemarijn T. Steensma
Remy M.J.P. Rikers
Axel P.N. Themmen
Jan N.M. IJzermans

Submitted

ABSTRACT

Purpose of the study: Selection of surgical residents is a high-stake process entailing the great responsibility to correctly identify and recruit successful candidates. However, information on factors that may affect performance during surgical residency training is limited. This study aimed to determine the relation between pre-training variables and indicators for performance during surgical training.

Study design: In 2014, application files of a cohort surgical residents enrolled in a regionally organized training program between 2007 and 2013 in the Netherlands were retrospectively analyzed. Pre-training variables included: demographic data, secondary and medical school grades, academic credentials and relevant clinical work experience. Using regression analyses, these variables were related to scores for in-training examinations completed between 2011-2013 and faculty evaluations obtained by telephone interview in August 2013.

Results: The mean (±SD) age at time of application of the 43 surgical residents (24 men) was 28.4 ±1.4 years. Twenty-seven percent (adjusted R2=0.266, F=8.418, p=0.001) of the variance for the in-training examinations was explained by gender (B=-8.130, p=0.005) and mean preclinical grade in medical school (B=10.475, p=0.002). For the faculty evaluations, 39% (adjusted R2=0.388, F=7.661, p<0.001) of the variance was assigned to: type of secondary school (B=-5.046, p=0.032), graduation grade for Mathematics (B=3.186, p=0.004), mean preclinical grade in medical school (B=5.327, p=0.044) and length of clinical work experience (B=-0.292, p=0.025).

Conclusions: Pre-training variables derived from structured curriculum vita as well as secondary and medical school diplomas were associated with indicators for performance during surgical training.

INTRODUCTION

The aim of surgical residency training is to adequately prepare the next generation of surgeons to independently perform the full scope of tasks they are expected to carry out in professional practice, and thus guarantee the continued delivery of high-quality patient care. In order to meet contemporary healthcare demands, such as a rising awareness for accountability, transparency and quality in medicine by an increasingly well-informed public, training surgeons that match today's practice needs is critical.[1] The goal for selection committees is therefore to ensure that candidates who acquire skills quickly and will work safe and competent in their future careers are attracted into surgical training programs, stressing the importance of valid selection methods.

The total costs of surgical training are high, ranging from around £413.000 in the United Kingdom (UK), about $650.000 in the United States (US), to €840.000 in the Netherlands per full-time equivalent. These large investments are either financed with public money, funded by insurance companies, or paid personally (using private loans).[2-4] Therefore, the recruitment process should be directed to select candidates with a low risk of failure to prevent financial loss. In addition, the use of public money stresses the importance of appropriately functioning training programs in order to meet societal demands and bear social responsibility. For these reasons, selection of surgical residents is a high-stake and challenging process entailing the great responsibility to correctly identify and recruit successful candidates. However, reported attrition rates in surgical training programs are high.[5,6] To minimize the social and financial costs caused by attrition, more evidence-based information about the predictive validity of selection methods is clearly needed to optimize the selection procedure. To date, there is a paucity of information about the factors known at the time of application that can be related to performance during surgical residency training or beyond.[7]

A typical selection procedure for surgical residency training starts with screening of the cover letters and curriculum vitae provided by the candidates. Subsequently, those considered promising are invited for a personal interview.[8] As admission to surgical training programs in most Western countries is highly competitive,[9] it is not uncommon for aspiring surgeons to undertake various activities to improve their eligibility. These include participation in scientific research, writing a PhD thesis, and acquiring clinical experience and technical skills by working as a surgical house officer (UK) or a non-designated preliminary surgical resident (US).

Criteria that selection committees traditionally rely on include: cover letters, curriculum vitae, personal statements, letters of recommendation and academic credentials. However, the predictive validity of these selection tools is inconclusive.[10] Furthermore, the value of the personal interview, which is generally considered to be the decisive factor for selection committees in assigning a training post,[8,11] has been called into question.[7] Taken together, more information about the factors known at the time of application is highly requested in order to ensure that the best candidates are attracted into a surgical training program. This study aimed to determine the relation between pre-training variables that were derived from structured curriculum vitae, as well as secondary and medical school diplomas and indicators for subsequent performance during surgical residency training as reflected by scores for in-training examinations and evaluations by faculty surgeons.

METHODS

Setting

In the Netherlands, training in general surgery lasts 6 years and is organized in 8 training regions, each consisting of 1 university hospital and several affiliated district hospitals. After graduating medical school, consisting of a 4-year preclinical period in which mainly basic science, human organ systems and pathophysiological processes are educated followed by a 2-year clinical phase of rotational clerkships in various medical disciplines, those having the ambition to pursue a career in general surgery are allowed to apply for a training post. Biannually, a nationwide selection round takes place that is overseen and coordinated by the Association of Surgeons of the Netherlands. Candidates have to specify 2 training regions of preference and are required to submit a cover letter and a structured curriculum vitae together with supporting documents such as copies of secondary and medical school diplomas. The selection committees of the preferred training regions, consisting of the local program directors of the hospitals involved in surgical training and a resident-representative, eventually decide which candidates are invited for a personal interview and which of them are finally selected.[12,13]

Study design

This study was conducted in a surgical training region in the Netherlands, consisting of 1 university hospital and 6 district hospitals. From April to May 2014, 2 of the authors independently retrospectively analyzed the submitted application files of a cohort of surgical residents enrolled in the training program between 2007 and 2013. The pre-training variables derived from these files were related to indicators for subsequent performance during training, including scores for in-training examinations completed between 2011-2013 and faculty evaluations obtained by telephone interviews by the program director in August 2013. Ethical approval for this study was obtained from the local ethical research committee.

Pre-training variables

The variables entered into an anonymized database consisted of demographic details, such as gender and age. Performance during secondary school, taking 6 years, was analyzed by extracting graduation grades obtained at the national exam for the courses Biology, Chemistry, Mathematics, Physics and the languages Dutch and English (all prerequisites to enter medical school). In addition, since in the Netherlands 2 separate types of secondary school allow admission to university education, the influence of both these different school types was measured. Briefly, type I secondary schools, also called Athenaeum, differ from the type II schools, which are called Gymnasium or Lyceum (continental Europe equivalent to British grammar schools or North American university-preparatory schools), in a way that students are educated in Greek and/or Latin as an additional course. Medical school performance was assessed by calculating the mean grade obtained for written knowledge tests during the preclinical period, the mean score for the clinical phase of rotational clerkships, the awarded score for the surgery clerkship in particular, and finally the grade obtained for graduation research. It should be noted that grades and test scores in the Netherlands range from 1-10, with 10 being the highest available score and a score of 5.5 being the pass/fail threshold. Academic credentials were reflected by the number of publications, both published and accepted, and the number of oral or poster presentations at national and international conferences. In addition, a completed PhD thesis at the time of application or within one year after starting surgical residency training was considered an academic achievement. Finally, the number of months being employed as a surgical house officer or non-designated preliminary surgical resident was evaluated and considered an indicator for relevant clinical work experience.

Performance indicators

The first indicator for performance during surgical residency training were scores for in-training examinations completed between 2011 and 2013. These computer-based tests take place annually and consist of 90 multiple-choice questions that address medical knowledge

and the management of surgical patients regarding diagnosis and treatment. Scores are reported as the percentage of the number of questions answered correctly. Participation in these examinations is highly recommended but is not obligatory. The second indicator for performance was a score of clinical performance. Different faculty surgeons were asked to indicate their trust in the ability for each of the surgical residents in their hospital to treat/operate on their relatives. Evaluations of the faculty surgeons were obtained by telephone interview in August 2013 with scores ranging from 0 (indicating no confidence) to 100 (full confidence).

Statistical analysis

The data were analyzed with IBM SPSS Statistics version 20.0 (IBM Corp., Armonk, NY, US). Descriptive statistics were calculated for all pre-training variables. Missing data (<5%) were replaced with substituted values using the imputation method. Scores for the in-training examinations were corrected for level of difficulty by calculating z-scores for each of the analyzed years separately. Subsequently, these z-scores were further corrected by adjusting for postgraduate year of training, thereby taking into account that surgical residents' level of knowledge develops gradually during training. For the evaluations by the faculty surgeons, the inter-rater reliability was computed, using intra class correlations (ICC), in order to determine the extent of agreement between different raters. A one-way random effects model of the ICC was used as individual surgical residents were evaluated by different faculty surgeons, and average measures were determined since these type of evaluations should be completed by multiple raters. A Pearson correlation coefficient was calculated to determine the relation between both the performance indicators used in this study. A stepwise regression analysis with a backward elimination approach was carried out to identify pre-training variables that were related to scores for the in-training examinations and the faculty evaluations. The level for statistical significance was set at $p<0.05$.

RESULTS

Demographics

In total, application files of 44 surgical residents were analyzed. As a consequence of missing data, 1 resident was excluded from further analysis. Of the 43 remaining, 24 (56%) were men and 19 (44%) women. The mean (±SD) age at the time of application was 28.4 ±1.4 years. Twenty-nine of these residents (67%) had a type I secondary school degree. An overview of the other pre-training variables can be found in Table 1.

Table 1. Descriptive statistics of the pre-training variables derived from the files submitted at the time of application (n=43 surgical residents).

		Mean	SD	Range
Secondary school	Biology	7.04	0.74	6.0 - 9.0
	Chemistry	6.98	1.05	5.0 - 9.0
	Mathematics	7.13	1.07	5.0 - 9.0
	Physics	6.96	1.07	5.0 - 9.0
	Dutch	7.12	0.73	6.0 - 9.0
	English	6.91	0.97	5.0 - 9.0
Medical school	Preclinical grade	6.92	0.44	6.2 - 8.5
	Clerkship grade	8.08	0.35	7.0 - 8.7
	Surgery clerkship	8.21	0.41	7.0 - 9.0
	Graduation research	8.40	0.78	6.9 – 10.0
Academic credentials	Published and accepted articles *(median)*	8		1 - 27
	Oral and/or poster presentations *(median)*	8		0-20
	Completed PhD thesis* *(percentage)*	15 (35%)		
Clinical work experience in months		11.8	8.6	0 - 36

Data are presented as means and standard deviations (SD), unless otherwise indicated

*completed at the time of application or within 1 year after starting surgical residency training

Correlation between performance indicators

The scores for the in-training examinations correlated modestly to the evaluations by the faculty surgeons (r=0.309, p=0.047), indicating that these evaluations most probably refer to other qualities than the ready availability of medical knowledge.

Variables related to in-training examination scores

The mean (±SD) score for the annual in-training examinations was 48% ±7.9 (range 32-62) in 2011, 55% ±6.8 (range 44-67) in 2012, and 49% ±6.2 (range 37-62) in 2013. The median number of in-training examinations performed per individual resident was 2 (range 1-3). Regression analyses revealed that 27% percent (adjusted R2 =0.266, F=8.418, p=0.001) of the variance was explained by gender (B=-8.130, p=0.005), with men outperforming women, and by the mean preclinical grade in medical school (B=10.475, p=0.002)

Variables related to faculty evaluations

The mean (±SD) assigned score for the evaluations by the faculty surgeons was 76 ±8.5 (range 57-93). Per individual resident, a median number of 5 (range 4-6) evaluations were completed. The ICC was 0.738, indicating substantial extent of agreement between the different raters. Thirty-nine percent (adjusted R2 =0.388, F=7.661, p<0.001) of the variance for the evaluations was assigned to type of secondary school (B=-5.046, p=0.032), with type I outperforming type II. In addition, the graduation grade for Mathematics in secondary school (B=3.186, p=0.004), mean preclinical grade in medical school (B=5.327, p=0.044), and length of clinical work experience (B =-0.292, p=0.025) were found to be related.

DISCUSSION

Acceptance into surgical residency training is still not based tangible indicators that are scientifically proven. Historically, selection is based on secondary school passing examinations, medical school grades, academic credentials and letters from acknowledged academicians, followed by a personal interview.[8] Findings of the studies previously examining how selection criteria and their predictive value impact on surgical performance revealed conflicting results,[14,15] and were of limited quality.[16-18] Using data from application files of a cohort of surgical residents, this study aimed to determine whether pre-training variables could be associated with indicators for performance during surgical training. Indicators for performance were both objectively and subjectively measured. The objective indicator included scores for in-training examinations in which surgical residents' medical knowledge was assessed. The subjective indicator comprised evaluations by different faculty surgeons regarding their trust to let their relatives be taken care for by a surgical resident employed in their hospital.

Of all pre-training variables, only the mean preclinical grade in medical school was related to both the in-training examination scores and the evaluations by the faculty surgeons. During the preclinical years of medical school, basic science courses, studies on human organ systems and pathophysiological processes are integrated in thematic blocks, providing a fundamental basis for clinical medicine. Accordingly, it may be assumed that grades obtained for written knowledge tests during this phase relate to performance in later years. This finding corresponds to evidence that medical students' examination scores appear to be the best predictive markers of performance during the later years of training.[7]

Interestingly, gender was observed as a pre-training variable related to the score for the in-training examinations. Although it has been shown that women perform better during medical school,[19] results of this study demonstrate that this finding cannot be translated to the years of surgical residency training, with men outperforming their female colleagues. One explanation could be that women have a different view on their career. It has been reported that women tend to be more focused on a good balance between work and private life, may prefer more flexible hours, and may be less committed or attracted to seek leadership or management roles, to do research or to teach.[20] In addition, it may be that female surgical residents differently focus their energy and are less competitive to score highly on the in-training examinations. Furthermore, it may be speculated that the overlap between the years of surgical training and the years of childbearing and parenting may play a role as well.

It is of interest to note that that the evaluations by the faculty surgeons did not show a gender association. Nonetheless, although still underrepresented, women are increasingly entering the surgical workforce.[21] Also in the present study this demographic change is observed, with 44% of the surgical residents being a woman.

The mean score for the faculty evaluations was related to the type of secondary school. It was found that surgical residents with a type I secondary school degree received better evaluations as compared to those with a type II secondary school degree. This was contrary to our expectations since it seems conceivable that familiarity with Greek or Latin (=type II secondary school) is beneficial for medical performance and may be more demanding. However, this study and previous research show no support for this presumption.[22] One could argue that students graduating from a type II secondary school are more linguistically, culturally and philosophically oriented, skills that may be less relevant for a technical-oriented specialty as surgery. From this perspective, it would be interesting to investigate whether residents with a type II secondary school degree enrolled in more-person oriented specialties, such as internal medicine and neurology, perform better.

The graduation grade for Mathematics in secondary school was found to be associated with the clinical performance of surgical residents, as reflected by the faculty evaluations. Mathematics relates to the degree one is able to think across levels of abstraction, while abstraction is fundamental for cognitive development. Many studies have found strong relations between secondary school grades and preclinical performance,[23] but these appear to decrease during the clerkships.[24] Unexpectedly, no relation could be found with the in-training examination scores, probably because these scores are more directed to assess declarative knowledge as opposed to procedural knowledge.

Surprisingly, the length of clinical work experience demonstrated to be unfavorably related to clinical performance during training. At first glance this is unforeseen as more experience is expected to relate to better clinical performance. It may be that candidates who are motivated to become a surgeon and considered 'ready for practice' by his/her local program director are encouraged to apply for a training position at an early stage. As a result, these better candidates are also the ones who are selected early compared to those who do not have everything planned and organized.

Some pre-training variables were unexpectedly not related to performance during surgical residency training. These variables included, the mean score for the clinical phase of rotational clerkships and more specifically the surgery clerkship grade. Further analysis of the data

revealed that these variables are probably less informative as a serious restriction of range was observed, with mean scores >8.0 and small standard deviations. This finding indicates that clerkship scores, which are generally awarded after an oral examination and thus are not formally evaluated with a written test, have a weak discriminative ability and are prone to subjectivity that might be tainted by the relationship between student and assessor. A previous study showed that clerkship scores acquired at different US medical schools are difficult to compare and are not a reliable indicator of performance.[25] In contrast to findings in a systematic review examining the predictive value of selection methods for surgical residency training,[26] academic credentials, such as the number of both publications and presentations and a completed PhD thesis, were not associated with performance during surgical residency training in this study.

A limitation of this study is that data of candidates that were selected to enter the surgical training program could not be compared to those that were not. Unfortunately, in the Netherlands it is mandatory by law that application files of candidates who have not been selected are destructed following a selection round. Other limitations include the relatively small number of surgical residents that could be analyzed and the fact that this study was limited to a regionally organized training program. Therefore, a nationwide study with a larger sample size is needed to confirm the generalizability of the results found in this study. Moreover, a greater number of study objects would also allow to analyze data of surgical residents that dropped out of training, potentially providing selection committees better insight in which pre-training variables should be taken into account when recruiting candidates to enter a surgical training program.

CONCLUSION

In conclusion, this study identified that the pre-training variables gender, type of secondary school, graduation grade for Mathematics, average pre-clinical grade in medical school and length of clinical work experience were associated with indicators for performance during surgical residency training. Using these variables could be helpful to optimize the current selection process of surgical residents and may reduce dropout rates in the future. However, in addition to the lack of a widely used measure of success in surgical training and a limited body of literature reporting about long-term performance outcomes in professional practice, more work is needed to determine whether the wide range of selection criteria commonly used are truly predictive for successful performance during surgical residency training in an evidence-based manner.

REFERENCES

1. Frenk J, Chen L, Bhutta ZA, Cohen J, Crisp N, Evans T, et al. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. Lancet 2010;376(9756): 1923-1958.
2. The Cost of Surgical Training. Position statement by the Association of Surgeons in Training. http://www.asit.org/assets/documents/ASiT_Cost_of_Surgical_Training_final.pdf May 2007].
3. Wynn BO, Robert Smalley and Kristina M. Cordasco. Does It Cost More to Train Residents or to Replace Them? A Look at the Costs and Benefits of Operating Graduate Medical Education Programs. Santa Monica, CA: RAND Corporation, 2013. http://www.rand.org/pubs/research_reports/RR324.
4. Kostprijsonderzoek tot Medisch Specialist door Samenwerkende Topklinische Ziekenhuizen www.stz.nl/download/f61e145f90ee8e4b94c496d40445eb541471.pdf November 2012].
5. Dodson TF, Webb AL. Why do residents leave general surgery? The hidden problem in today's programs. Curr Surg 2005;62(1): 128-131.
6. Longo WE, Seashore J, Duffy A, Udelsman R. Attrition of categoric general surgery residents: results of a 20-year audit. Am J Surg 2009;197(6): 774-778; discussion 779-780.
7. Kenny S, McInnes M, Singh V. Associations between residency selection strategies and doctor performance: a meta-analysis. Medical Education 2013;47(8): 790-800.
8. Makdisi G, Takeuchi T, Rodriguez J, Rucinski J, Wise L. How We Select Our Residents—A Survey of Selection Criteria in General Surgery Residents. Journal of Surgical Education 2011;68(1): 67-72.
9. Carr AS, Munsch C, Buggle S, Hamilton P. Core surgical training and progression into specialty surgical training: how do we get the balance right? Bulletin of The Royal College of Surgeons of England 2011;93(7): 244-248.
10. McGaghie WC, Cohen ER, Wayne DB. Are United States Medical Licensing Exam Step 1 and 2 scores valid measures for postgraduate medical residency selection decisions? Academic medicine : journal of the Association of American Medical Colleges 2011;86(1): 48-52.
11. Melendez MM, Xu X, Sexton TR, Shapiro MJ, Mohan EP. The Importance of Basic Science and Clinical Research as a Selection Criterion for General Surgery Residency Programs. Journal of Surgical Education 2008;65(2): 151-154.
12. Bonjer HJ, Bruining HA. Surgery in the Netherlands. Arch Surg 1999;134(1): 92-98.
13. Borel-Rinkes IH, Gouma DJ, Hamming JF. Surgical training in the Netherlands. World J Surg 2008;32(10): 2172-2177.
14. Sutton E, Richardson JD, Ziegler C, Bond J, Burke-Poole M, McMasters KM. Is USMLE Step 1 score a valid predictor of success in surgical residency? Am J Surg 2014;208(6): 1029-1034.

15. Thordarson DB, Ebramzadeh E, Sangiorgio SN, Schnall SB, Patzakis MJ. Resident selection: How we are doing and why? Clinical orthopaedics and related research 2007(459): 255-259.
16. Oldfield Z, Beasley SW, Smith J, Anthony A, Watt A. Correlation of selection scores with subsequent assessment scores during surgical training. ANZ J Surg 2013;83(6): 412-416.
17. Grewal SG, Yeung LS, Brandes SB. Predictors of success in a urology residency program. J Surg Educ 2013;70(1): 138-143.
18. Fryer JP, Corcoran N, George B, Wang E, Darosa D. Does resident ranking during recruitment accurately predict subsequent performance as a surgical resident? J Surg Educ 2012;69(6): 724-730.
19. Haq I, Higham J, Morris R, Dacre J. Effect of ethnicity and gender on performance in undergraduate medical examinations. Medical Education 2005;39(11): 1126-1128.
20. Sanfey HA, Fromson JA, Mellinger JD, Rakinic J, Williams M, Williams B. Surgeons in difficulty: an exploration of behavior differences among male and female surgeons. Journal of the American College of Surgeons 2014;219(4): e141.
21. Roupret M, Maggiori L, Lefevre JH. Upcoming female preponderance within surgery residents and the association of sex with the surgical career choice in the new millennium: results from a national survey in France. Am J Surg 2011;202(2): 237-242.
22. Pampush JD, Petto AJ. Familiarity with Latin and Greek anatomical terms and course performance in undergraduates. Anatomical Sciences Education 2011;4(1): 9-15.
23. Ferguson E, James D, Madeley L. Factors associated with success in medical school: systematic review of the literature. Bmj 2002;324(7343): 952-957.
24. McManus IC, Smithers E, Partridge P, Keeling A, Fleming PR. A levels and intelligence as predictors of medical careers in UK doctors: 20 year prospective study. Bmj 2003;327(7407): 139-142.
25. Takayama H, Grinsell R, Brock D, Foy H, Pellegrini C, Horvath K. Is it Appropriate to Use Core Clerkship Grades in the Selection of Residents? Current Surgery 2006;63(6): 391-396.
26. Maan ZN, Maan IN, Darzi AW, Aggarwal R. Systematic review of predictors of surgical performance. Br J Surg 2012;99(12): 1610-1621.

Chapter 8

Summary and conclusions
Nederlandse samenvatting en conclusies

Cornelis J. Hopmans

SUMMARY AND CONCLUSIONS

Within the past decade, the structure and format of surgical residency training has changed radically by the introduction of competency-based training programs, the progressive fragmentation of general surgery into subspecialties, and the implementation of stringent work hour restrictions. The aim of this thesis was to explore how these developments have transformed the landscape of surgical residency training within the past years. The research described in this thesis has been conducted in a surgical training region located in the Southwest of the Netherlands, consisting of 1 university hospital and 6 affiliated district hospitals. Conclusions resulting from this research may be helpful for stakeholders responsible for the organization of training in general surgery in the Netherlands and could be used for continuous improvement of the surgical residency training program and to guide future directions in surgical training, hence ensuring high-quality training and well-balanced education for the next generation of surgeons.

In **chapter 2** the importance of the various Canadian Medical Education Directives for Specialists (CanMEDS) competencies was investigated, using a survey distributed among surgical residents (n=51) and attending surgeons (n=108). In addition, the suitability of the adopted assessment tools to evaluate and monitor competence progression was surveyed. With a response rate of 88%, the results showed that surgical residents and attending surgeons did not recognize that all CanMEDS competencies are equally important for the competency profile of a surgeon, undervaluing the general competencies Manager and Health Advocate. The competency Medical Expert was considered the most important. The assessment tools were predominantly considered unsuitable for competence evaluation. An exception were the instruments In-Training Evaluation Report (ITER) and Objective Structured Assessment of Technical Skills (OSATS). These findings support the traditional surgical point of view that it is paramount to have a broad base of medical knowledge and confirm that personal interaction with attending surgeons and evidence for operative skills are still considered the foundation for surgical residency training.

In **chapter 3** the reform of surgical residency training in Netherlands was examined, using a mixed-methods study design. The focus of research was to examine how surgical residents and attending surgeons have implemented and incorporated a competency-based training and assessment program, how external policy reforms are given shape and how this has affected training in daily clinical practice. Data were derived from semi-structured interviews

(n=26), observations during meetings of an expert group (n=7), and a questionnaire (response 140/159). The results of this study revealed that both surgical residents and attending surgeons are willingly to embrace the modern training methodology, with its underlying competency framework, assessment tools and subspecialty training. However, at the same time, a sense of skepticism and unease was observed, especially regarding the work hour restrictions. To respond and deal with the sweeping transformation of surgical residency training, so called 'repair work' is conducted. This goes beyond restoring the traditional way of surgical residency training in the sense of re-establishing the past. Instead, it is an interactive process: implementing the imposed training methodology and adapting to external policy reforms, while simultaneously seeking convenient and feasible solutions to preserve certain traditional values and related practices.

Within competency-based curricula, the instrument Objective Structured Assessment of Technical Skills (OSATS) is considered the standard for the assessment and evaluation of operative skills. To rely on this instrument and to defend its widespread implementation, it is imperative to have sufficient evidence comprising metrics for validity and reliability. Although studies in skills laboratory settings on bench model simulations, as well as in animal models and cadaveric training models have demonstrated evidence for the usefulness of OSATS, data supporting application of this instrument in the operating room are limited. **Chapter 4** presents the result of a prospective observational study aimed at evaluating the validity and reliability of the OSATS instrument to assess surgical residents' operative skills in an operating theater setting. The participants in this study were selected from the 7 hospitals within this surgical training region and classified equally into 3 groups according to postgraduate training year (group 1: PGY 1-2, group 2: PGY 3-4, group 3 PGY: 5-6). Each participant had to perform 5 different types of surgical procedures. Surgical performance was measured using a modified OSATS instrument consisting of 3 scales: Global Rating Scale, Overall Performance Scale and Alphabetic Summary Scale. Comparisons of the performance scores on the different rating scales generally demonstrated improvement along with the training level, which constitutes evidence for construct validity. Evidence for reliability was shown by the high internal consistency reliability coefficients for scores on the Global Rating Scale. Given its weak discriminative ability, use of the Overall Performance Scale should be reconsidered. The Alphabetic Summary Scale, on the other hand, seems to be of additional value for clinical practice. Nevertheless, it must be pointed out that, despite observational instruments such as OSATS may currently be the most accepted and easily available assessment method to evaluate surgical residents' operative skills in the operating room, performance scores on these instruments are inherently associated with subjectivity and do not necessarily provide an accurate and objective indication for operative competence. Finally, it should be noted that

a modified version of the original OSATS concept has been examined in this study, without scientific evidence supporting these modifications.

In **chapter 5** it was examined whether the progressive reduction of the working week to 48 hours resulting from implementation of the European Working Time Directive (EWTD) has affected the operative experience of surgical residents. The data for this retrospective analysis were derived from operating records in which the surgical procedures conducted in the operating theaters of the hospitals within this training region for the period 2005-2012 are summarized. The results of this study revealed that implementation of the EWTD has not adversely affected the number of operations performed by surgical residents, indicating that operative experience has been preserved. More efficient use of available training time in combination with an increased involvement of healthcare workers for routine surgical work on the wards and outpatient clinic may be an explanation for this outcome. On the other hand, it could also imply that with the implementation of the EWTD there is an increased focus on the preservation of operative skills, while attention paid to other important training aspects, such as attendance during clinical handovers, multidisciplinary meetings or teaching sessions, is compromised. In addition, it might be that surgical residents are less exposed to critical phases in the management of patients, such as during time of admission, decision making in the intensive care unit, grand rounds on the inpatient wards, or follow-up at the outpatient clinic.

Over the past decade, surgical training has changed by the introduction of competency-based curricula, implementation of work hour restrictions, and the fragmentation of general surgery into subspecialties. In **chapter 6**, it was investigated whether the attitude of surgical residents and attending surgeons toward these developments has changed over time. Results of this study are based on a survey that was distributed in 2011 and 2013, yielding a response rate of 88% (140/159) in 2011 and 82% (127/154) in 2013. The main findings of this study show that both surgical residents and attending surgeons prefer the traditional master-apprenticeship training model over the competency-based curriculum, whereas the importance of the CanMEDS competencies is recognized. The adopted assessment tools are predominantly considered unsuitable to evaluate competence progression. Attending surgeons are significantly more concerned about the consequences of work hour restrictions on surgical training and healthcare delivery as compared to surgical residents. On completion of the training program, there is a feeling of unpreparedness for independent practice as surgical specialist in any of the subspecialties offered. The future of general surgery appears to be ill-defined. In conclusion, this interim analysis demonstrates that 4 years after surgical residency training has radically been changed, contemporary developments evoke a desire to preserve traditional values and related practices, underlining the relevance of proper feedback to adjust change. The outcomes of the 6-year follow-up have to be awaited to define more precisely whether the attitude toward the transformation of surgical residency training has evolved over the years .

In the final chapter of this thesis, **chapter 7**, application files of a cohort surgical residents enrolled in the surgical training program between 2007 and 2013 were retrospectively analyzed. From these files, demographic data, secondary and medical school grades, academic credentials and relevant clinical work experience were derived. These pre-training variables were related to scores for in-training examinations completed between 2011-2013 and faculty evaluations obtained by telephone interview in August 2013. The results of this study showed that the mean preclinical grade in medical school was related to both the in-training examination scores and the faculty evaluation scores. Gender was found to be associated with scores for the in-training examinations, with men outperforming their female colleagues. The faculty evaluation scores could be related to the type of secondary school (with type I = Athenaeum outperforming type II = Gymnasium/ Lyceum) and the graduation grade for Mathematics in secondary school. The length of clinical working experience demonstrated to be unfavorably related to the faculty evaluation scores. Attention for these variables could be helpful to optimize the current selection process of surgical residents and potentially reduce dropout rates. However, in addition to the lack of a widely used measure of success in surgical training and a limited body of literature reporting about long-term performance outcomes in professional practice, more work is needed to determine whether the wide range of selection criteria commonly used are truly predictive for successful performance during surgical residency training. Finally, it should be noted that non-cognitive skills, such as motivation, perseverance and self-discipline, may play a critical role in educational outcomes and success in professional practice.

NEDERLANDSE SAMENVATTING EN CONCLUSIES

In het afgelopen decennium is de structuur van de opleiding Heelkunde drastisch veranderd door de introductie van een competentiegericht opleidingsplan, de opsplitsing van de algemene chirurgie in aandachtsgebieden en de stapsgewijze reductie van de gemiddelde arbeidsweek voor artsen in opleiding tot medisch specialist (AIOS) naar 48 uur. In dit proefschrift is door middel van verschillende studies onderzocht hoe deze ontwikkelingen het opleidingslandschap in de afgelopen jaren hebben veranderd. De studies in dit proefschrift zijn uitgevoerd in Onderwijs- en Opleidingsregio Rotterdam (regio IV). De betrokken ziekenhuizen zijn het Erasmus MC en de 6 affiliatie ziekenhuizen (Amphia Ziekenhuis, Breda; Ikazia Ziekenhuis, Rotterdam; Maasstad Ziekenhuis, Rotterdam; Reinier de Graaf Groep, Delft; Franciscus Gasthuis & Vlietland, Rotterdam; IJsselland Ziekenhuis, Capelle aan den IJssel) die samen de opleiding Heelkunde in deze regio verzorgen. De conclusies voorkomend uit deze studies kunnen van waarde zijn voor stakeholders die verantwoordelijk zijn voor de structuur en organisatie van de opleiding Heelkunde en kunnen mogelijk een rol spelen bij het continue verbeteringsproces van het opleidingsplan waardoor een hoogwaardige opleiding en dito onderwijs voor toekomstig chirurgen blijft gewaarborgd.

In **hoofdstuk 2** worden de resultaten van een vragenlijst besproken waarin 51 AIOS en 108 chirurgen uit de Onderwijs- en Opleidingsregio Rotterdam werden gevraagd om de waarde van de verschillende CanMEDS competenties voor het competentieprofiel van een chirurg aan te geven. Daarnaast werd gevraagd om te beoordelen hoe geschikt de verschillende toetsinstrumenten zijn om verwezenlijking van de CanMEDS competenties te meten. Het totaal aantal geretourneerde vragenlijsten bedroeg 140/159 (88%). De resultaten laten zien dat niet alle CanMEDS competenties op vergelijkbare wijze bijdragen aan het competentieprofiel van een chirurg. De competentie Medisch Handelen wordt als meest belangrijk gescoord, terwijl de generieke competenties Maatschappelijk Handelen en Organisatie ondergewaardeerd worden. De toetsinstrumenten worden overwegend als ongeschikt ervaren om realisering van de CanMEDS competenties te beoordelen. Uitzonderingen hierop zijn het periodiek voortgangsgesprek en het instrument OSATS dat gebruikt wordt om de operatieve vaardigheden van AIOS in de operatiekamer te meten. Deze bevindingen onderschrijven het traditioneel chirurgisch standpunt dat gedegen medische kennis van zeer grote waarde is en bevestigen dat persoonlijk contact tussen AIOS en chirurg in combinatie met bewijs voor chirurgische vaardigheid nog steeds het fundament voor de opleiding Heelkunde vormt.

In **hoofdstuk 3** wordt een mixed-methods studie gepresenteerd. Dit houdt in dat zowel kwalitatieve als kwantitatieve onderzoeksmethoden zijn gebruikt om te analyseren hoe de implementatie van een competentiegericht opleidingsplan in Onderwijs- en Opleidingsregio Rotterdam wordt ervaren en hoe er wordt aangekeken tegen werktijdenreductie en differentiatie van de algemene chirurgie. De data werden verkregen uit semi-gestructureerde interviews (n=26), observaties tijdens bijeenkomsten van de regionale opleidingscommissie (n=7) en een vragenlijst (respons 140/159). De resultaten van het onderzoek laten zien dat zowel AIOS als chirurgen zich welwillend opstellen tegenover de moderne opleidingsvorm waarin aandacht is voor competentieontwikkeling, gestructureerde momenten van toetsing en feedback en subspecialisatie in een aandachtsgebied. Tegelijkertijd bleek er ook een gevoel van scepsis en onbehagen te bestaan, met name ten aanzien van de maximale duur van een werkweek die vastgesteld is op 48 uur. Om in de dagelijkse praktijk in te spelen op deze ingrijpende metamorfose van de opleiding Heelkunde, wordt er zogeheten 'repair work' verricht. Dit betekent niet dat er naar wordt gestreefd om vast te houden aan de traditionele manier van opleiden. In plaats daarvan wordt er getracht om de opgelegde curriculumherziening tot een succes te maken en tegemoet te komen aan beleidshervormingen, terwijl er tegelijkertijd gezocht wordt naar manieren om bepaalde tradities en waarden uit het verleden te behouden.

In **hoofdstuk 4** wordt de validiteit en betrouwbaarheid van het toetsinstrument OSATS geanalyseerd en besproken. Dit instrument wordt gebruikt om de operatieve vaardigheden van AIOS te beoordelen en inzichtelijk te maken. Bewijs voor het nut van dit instrument is voornamelijk aangetoond in experimentele studies in skillslaboratoria maar in de dagelijkse praktijk slechts zeer beperkt onderzocht. Voor dit prospectief observationeel onderzoek zijn 24 AIOS uit de verschillende opleidingsziekenhuizen in de Onderwijs- en Opleidingsregio Rotterdam geselecteerd die evenredig werden verdeeld over 3 groepen op basis van opleidingsjaar (groep 1, opleidingsjaar 1-2; groep 2 opleidingsjaar 3-4; groep 3, opleidingsjaar 5-6). Iedere deelnemer werd gevraagd om 5 verschillende type operaties 6 keer uit te voeren onder supervisie van een chirurg. Deze operaties werden vervolgens geëvalueerd met een gemodificeerde versie van het originele OSATS instrument bestaande uit 3 verschillende beoordelingsschalen: een algemene beoordelingsschaal, een numeriek eindoordeel en een alfabetische prestatieschaal. Vergelijkingen van de scores op de verschillende beoordelingsschalen lieten over het algemeen zien dat met het opleidingsniveau van de AIOS ook de score toenam. Dit toont bewijs voor constructvaliditeit aan. Bewijs voor betrouwbaarheid werd geleverd door de hoge interne –consistentie coëfficiënten voor scores op de algemene beoordelingsschaal. Gezien het zwak onderscheidend vermogen dient het gebruik van het numeriek eindoordeel te worden heroverwogen. De alfabetische prestatieschaal lijkt

daarentegen voor de klinische praktijk van toegevoegde waarde te zijn. Echter, ondanks dat het instrument OSATS momenteel de meest geaccepteerde en eenvoudigste beoordelingsmethode is om operatieve vaardigheden van AIOS in de operatiekamer te meten, moet er op worden gewezen dat beoordelingen op dit toetsinstrument inherent verbonden zijn aan subjectiviteit en niet noodzakelijkerwijs een accurate en objectieve indicatie zijn voor chirurgische bekwaamheid. Tenslotte moet worden opgemerkt dat in deze studie een gemodificeerde versie van het originele OSATS concept is onderzocht zonder dat er voor deze modificaties wetenschappelijke onderbouwing bestaat.

Hoofdstuk 5 beschrijft een retrospectieve studie waarin is onderzocht of de implementatie van het arbeidstijdenbesluit het aantal operaties dat door AIOS wordt verricht heeft beïnvloed. De data voor dit onderzoek werden verkregen uit visitatierapporten van de verschillende opleidingsziekenhuizen in de Onderwijs- en Opleidingsregio Rotterdam over het tijdvak 2005-2012. De conclusie van dit onderzoek is dat voortschrijdende werktijdenreductie geen invloed heeft gehad op het aantal operaties dat door AIOS wordt verricht. Efficiënter gebruik van de beschikbare opleidingstijd in combinatie met een toegenomen aantal specialistisch verpleegkundigen voor routinematig chirurgische werkzaamheden op de afdelingen en polikliniek zijn factoren die een verklaring voor deze uitkomst kunnen zijn. Het kan anderzijds ook betekenen dat er door de implementatie van het arbeidstijdenbesluit een toegenomen aandacht is voor het voor het behoud van operatieve vaardigheden, terwijl de nadruk op andere belangrijke aspecten van de opleiding in het gedrang komt. Hierbij moet worden gedacht aan opleidingsmomenten tijdens overdrachten, multidisciplinaire besprekingen en onderwijsbijeenkomsten. Daarnaast kan het zijn dat AIOS minder nauw betrokken zijn bij belangrijke momenten in het zorgproces, zoals tijdens het moment van opname, besluitvorming op de intensive care unit, visite op de afdeling of follow-up op de polikliniek.

In het afgelopen decennium is de structuur van de opleiding Heelkunde drastisch veranderd door de introductie van een competentiegericht opleidingsplan, de opsplitsing van de algemene chirurgie in aandachtsgebieden en de stapsgewijze reductie van de gemiddelde arbeidsweek voor artsen in opleiding tot medisch specialist (AIOS) naar 48 uur. **Hoofdstuk 6** beschrijft de resultaten van een onderzoek die zijn gebaseerd op een vragenlijst die in 2011 en 2013 werd verspreid waarin respondenten over bovenstaande hervormingen en ontwikkelingen werden bevraagd . Het responspercentage was op beide meetmomenten hoog en bedroeg respectievelijk 88% (140/159) in 2011 en 82% (127/154) in 2013. De belangrijkste bevindingen in deze studie zijn dat zowel AIOS en chirurgen het traditionele meester-gezel model prefereerden boven het competentiegerichte opleidingsplan. De verschillende CanMEDS competenties worden als belangrijk ervaren voor het competentieprofiel van een

chirurg. De toetsinstrumenten worden overwegend als ongeschikt beoordeeld om voortgang van de AIOS te meten. Chirurgen waren duidelijk meer bezorgd dan de AIOS over de potentiele gevolgen van het arbeidstijdenbesluit op de kwaliteit van de Heelkunde opleiding en de continuïteit van zorg. Na het afronden van de opleiding heerst er onder AIOS en chirurgen een gevoel onvoldoende voorbereid te zijn op de praktijk als gedifferentieerd chirurg. De toekomst van de Algemene Heelkunde is onzeker. Samengevat heeft deze interim analyse laten zien dat 4 jaar na de hervorming van de opleiding Heelkunde er een sterk verlangen bestaat om bepaalde tradities uit het meester-gezel tijdperk te behouden. Dit benadrukt de relevantie van feedback. Uit de resultaten van de herhaalmeting, die later dit jaar verricht zal worden, kan pas worden vastgesteld hoe men de recente ontwikkelingen die het opleidingslandschap in de afgelopen jaren ingrijpend hebben veranderd nu gewaardeerd worden.

Het laatste hoofdstuk van dit proefschrift, **hoofdstuk 7**, bespreekt de resultaten van een studie gericht op het identificeren van factoren die voorspellend zijn voor het functioneren tijdens de opleiding Heelkunde. De sollicitatiedossiers van 44 AIOS Heelkunde, die in de periode 2007-2013 in Onderwijs- en Opleidingsregio Rotterdam hun opleiding volgden, werden retrospectief geanalyseerd. Uit deze dossiers werden variabelen geëxtraheerd die vermeld staan in het curriculum vitae zoals biografische gegevens, klinische werkervaring en presaties op wetenschappelijk gebied. Daarnaast werden de prestaties van de sollicitanten op de middelbare school en tijdens de studie geneeskunde (propedeuse, doctoraal- en artsexamen) geïncludeerd als variabele. Deze gegevens werden gecorreleerd aan scores die behaald zijn tijdens een jaarlijkse kennistoets (periode 2011-2013) en uitkomstmaten die verkregen zijn uit telefonische interviews (augustus 2013) waarin verschillende chirurgen gevraagd werd om het algemeen medisch handelen van AIOS die in dit onderzoek geïncludeerd zijn te beoordelen. De resultaten laten zien dat er een positieve relatie bestaat tussen het gemiddelde cijfer behaald in de doctoraalfase tijdens de studie geneeskunde en de scores voor zowel de jaarlijkse kennistoets als de beoordeling voor het algemeen medisch handelen. Het mannelijk geslacht bleek geassocieerd te zijn met hogere scores voor de kennistoets. De uitkomstmaten voor het algemeen medisch handelen konden gerelateerd worden aan het type middelbare school (atheneumleerlingen presteerden beter dan gymnasiasten) en het eindexamen cijfer voor wiskunde op de middelbare school. Er bleek een negatieve relatie te bestaan tussen de duur van de klinische werkervaring en de score voor het algemeen medisch handelen. Aandacht voor deze variabelen zou nuttig kunnen zijn om het huidige sollicitatieproces te optimaliseren en in de toekomst uitval te verminderen. Gezien het ontbreken van een uitkomstmaat die succes tijdens de opleiding weerspiegelt en de zeer beperkte hoeveelheid literatuur waarin gerapporteerd wordt over de langetermijnprestaties tijdens de vervolgcarrière, is er meer onderzoek nodig om vast te stellen of het brede scala aan veelgebruikte selectiecriteria

daadwerkelijk een voorspellende waarde heeft voor het succesvol doorlopen van de opleiding Heelkunde. Tenslotte is het mogelijk dat niet-cognitieve persoonlijkheidskenmerken zoals motivatie, doorzettingsvermogen en zelfdiscipline een essentiële rol spelen bij zowel prestaties tijdens de opleiding als in de beroepspraktijk.

Chapter 9

Future perspectives

Cornelis J. Hopmans

COMPETENCY-BASED TRAINING

In most Western countries, the educational concept of surgical residency training has shifted from the traditional master-apprenticeship model toward competency-based curricula.[1,2] The key characteristic of competency-based training is that surgical residents only can obtain their board certification when competence has been demonstrated in the domains deemed necessary for professional practice. This means that not the duration of training is crucial, but the particular surgical skills or (sub)specialty for which one wants to qualify should be the decisive factor, allowing a flexible duration of training instead of a time-based training period.[3] Well-defined learning outcomes and more importantly assessment tools, which can confirm whether a certain level of competence has been achieved, are highly critical to successfully implement this novel approach to surgical training.[4]

Although the educational principle of competency-based training at first glance seems attractive, it has raised a number of issues and challenges. First, an important aspect of competency-based medical education is that, in addition to medical-technical skills, development of competencies that reflect general healthcare play a pivotal role. Effectively working within a team, adequate communication and information sharing with everyone involved in the patient's care, ensuring patient safety, attention for quality improvement, as well as respect for the patient's perspective in decision making processes are some examples of these general competencies.[5] However, recent studies, including our survey analysis (Chapter 2), have demonstrated that general competencies are not recognized as equally important as medical-technical skills.[6,7] To give general competencies their essential content and form, we advocate to conduct research to explore why they are not so popular. Subsequently, teaching strategies in how general competencies best can be integrated and evaluated in daily clinical practice are essential and should be either improved or developed. After all, it has been shown that adverse events in healthcare, and specifically surgery, are more likely to originate from inadequate professional competencies than a lack of medical knowledge or technical skills.[8,9]

Furthermore, the current literature provides no high-level evidence that the abundance of assessment tools developed are capable of providing an accurate evaluation of surgical residents' skills.[10,11] Instead, assessment tools are blamed for causing an administrative burden. In addition, it has been reported that a tick box perception prevails and personal preferences may affect evaluations, making that assessment tools are inherently associated with subjectivity.[12,13] Considering this, in combination with the fact that conducting high-quality studies

with the objective to evaluate assessment tools in a complex setting like a hospital is an extremely difficult undertaking, the debate regarding the role assessment tools should fulfill in clinical practice cannot be closed yet. In the Netherlands, entrustable professional activities (EPA's), which comprise clinical skills and abilities as well as more general facets of competence, have recently been incorporated in the surgical training program.[14] If the concept of EPA's can serve as a valid and reliable alternative evaluation method remains to be seen and should be focus of research.

WORK HOUR RESTRICTIONS

Several studies have demonstrated that surgical residents' personal well-being and lifestyle have improved since the implementation of work hour restrictions.[15,16] However, within the surgical profession, the implementation of work hour restrictions has led to widespread concern regarding the effects of decreased working hours on the quality of training and continuity of patient care.[17,18] In the past years, the impact of work hour restrictions on the quality of training has been extensively examined, showing that the caseload surgical residents are exposed to during their training has not adversely been affected.[19,20] However, in this context it should be noted that case numbers alone are not the surrogate of a well-trained surgeon. Therefore, operative volume might be an insufficient variable to measure the impact of work hour restrictions on the quality of surgical residency training. More studies are needed to establish the effect of reduced working hours on the development of other competencies that are essential for a surgeon, in addition to surgical skills. In addition, the impact of work hour restrictions on the continuity of patient care remains poorly understood with studies reporting conflicting results and theories.[21-24] Consequently, discussions on the effect of work hour restrictions are ongoing and future studies will remain necessary.

SUBSPECIALTY TRAINING

The increased demand of patients to be treated by specialized experts, innovative medical technologies, as well as the strong body of evidence showing a relationship between hospital volume and patient outcomes for a great number of surgical procedures, have fragmented general surgery into various subspecialties.[25,26] These subspecialties are transforming progressively from areas of interest within general surgery into recognized mono-specialties with their own training program and standards.[27,28] Whether training in general surgery in the future should be continued, remodeled or will disappear is a controversial topic in literature with legitimate pros and cons on either side of the debate. Opinions may vary greatly depending on various perspectives, such as geographic considerations (rural vs. urban) or practice desires (generalist vs. specialist).[25,29,30] For the densely populated Netherlands with modern infrastructure facilities, we recommend to develop a surgical training program that is aimed at subspecialty training in well-defined focus areas and suggest that remodeling of the current training schedule should be considered. After 2 years of common trunk, a period shared with other surgical specialties aimed at teaching the overarching principles of medical practice and basic operative skills, surgical residents should spend the final phase of training on subspecialty training in a focus area. As a consequence, investment in detailed medical knowledge, as well as complex surgical skills and techniques that have no relevance for the intended practice, will be avoided and would potentially allow reduction of training time. Completion of such a training pathway would produce surgeons that are adequately prepared for the challenging task to work in an increasing complex healthcare environment in which surgical care is practiced in a multidisciplinary setting with competing demands and the doctor-patient relationship is likely to transform into a business partnership. It is important to note that such a design of surgical residency training necessitates adaptation of healthcare organization. A smaller number of hospitals will be equipped to offer the full repertoire of surgical procedures. A national and regional strategy will be needed to guarantee access to healthcare facilities within a certain time and distance, during daytime as well as during out-of-office hours and weekends. To realize such a challenging and innovative healthcare landscape, a multitask approach is a prerequisite. All stakeholders, including government, health insurance companies, patient organizations, professional associations, as well as healthcare professionals, have a shared responsibility in this respect and must feel committed to contribute to the ongoing discussions. Obviously, the surgical profession must play a pivotal role and define the balance between organizational efficiency and patients' interests, realizing that the quality demanded today will define the standard of care in the future.

REFERENCES

1. Collins JP, Civil ID, Sugrue M, Balogh Z, Chehade MJ. Surgical education and training in Australia and New Zealand. World Journal of Surgery 2008;32(10): 2138-2144.
2. Smith AJ, Aggarwal R, Warren OJ, Paraskeva P. Surgical training and certification in the United kingdom. World J Surg 2009;33(2): 174-179.
3. Long DM. Competency-based residency training: the next advance in graduate medical education. Acad Med 2000;75(12): 1178-1183.
4. Epstein RM. Assessment in medical education. N Engl J Med 2007;356(4): 387-396.
5. Frank JR, Langer B. Collaboration, communication, management, and advocacy: teaching surgeons new skills through the CanMEDS Project. World J Surg 2003;27(8): 972-978; discussion 978.
6. Hopmans CJ, den Hoed PT, Wallenburg I, van der Laan L, van der Harst E, van der Elst M, et al. Surgeons' attitude toward a competency-based training and assessment program: results of a multicenter survey. J Surg Educ 2013;70(5): 647-654.
7. Arora S, Sevdalis N, Suliman I, Athanasiou T, Kneebone R, Darzi A. What makes a competent surgeon?: experts' and trainees' perceptions of the roles of a surgeon. Am J Surg 2009;198(5): 726-732.
8. Kripalani S, LeFevre F, Phillips CO, Williams MV, Basaviah P, Baker DW. Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care. JAMA 2007;297(8): 831-841.
9. Yule S, Paterson-Brown S. Surgeons' non-technical skills. Surg Clin North Am 2012;92(1): 37-50.
10. Carraccio C, Wolfsthal SD, Englander R, Ferentz K, Martin C. Shifting paradigms: from Flexner to competencies. Acad Med 2002;77(5): 361-367.
11. Lurie SJ, Mooney CJ, Lyness JM. Measurement of the general competencies of the accreditation council for graduate medical education: a systematic review. Acad Med 2009;84(3): 301-309.
12. Hopmans CJ, den Hoed PT, van der Laan L, van der Harst E, van der Elst M, Mannaerts GHH, et al. Assessment of surgery residents' operative skills in the operating theater using a modified Objective Structured Assessment of Technical Skills (OSATS): A prospective multicenter study. Surgery 2014.
13. Williams RG, Dunnington GL. Prognostic value of resident clinical performance ratings. J Am Coll Surg 2004;199(4): 620-627.
14. Dewey CM, Jonker G, Ten Cate O, Turner TL. Entrustable professional activities (EPAs) for teachers in medical education: Has the time come? Med Teach 2016: 1-3.
15. Curet MJ. Resident work hour restrictions: where are we now? J Am Coll Surg 2008;207(5): 767-776.
16. Jamal MH, Rousseau MC, Hanna WC, Doi SA, Meterissian S, Snell L. Effect of the ACGME duty hours restrictions on surgical residents and faculty: a systematic review. Acad Med 2011;86(1): 34-42.
17. Rosenbaum L, Lamas D. Residents' duty hours--toward an empirical narrative. N Engl J Med 2012;367(21): 2044-2049.

18. Roy P. Working time regulations for trainee doctors. BMJ 2009;339.
19. Hopmans CJ, den Hoed PT, van der Laan L, van der Harst E, van der Elst M, Mannearts GH, et al. The Operative Experience of Surgical Residents on the Other side of the Atlantic in the Era of Working Time Restrictions. Journal of the American College of Surgeons 2014;219(3, Supplement): S119.
20. Moonesinghe SR, Lowery J, Shahi N, Millen A, Beard JD. Impact of reduction in working hours for doctors in training on postgraduate medical education and patients' outcomes: systematic review. BMJ 2011;342: d1580.
21. Axelrod L, Shah DJ, Jena AB. The European Working Time Directive: an uncontrolled experiment in medical care and education. JAMA 2013;309(5): 447-448.
22. Barger LK, Ayas NT, Cade BE, Cronin JW, Rosner B, Speizer FE, et al. Impact of extended-duration shifts on medical errors, adverse events, and attentional failures. PLoS Med 2006;3(12): e487.
23. Cairns H, Hendry B, Leather A, Moxham J. Outcomes of the European Working Time Directive. BMJ 2008;337: a942.
24. Landrigan CP, Rothschild JM, Cronin JW, Kaushal R, Burdick E, Katz JT, et al. Effect of reducing interns' work hours on serious medical errors in intensive care units. N Engl J Med 2004;351(18): 1838-1848.
25. Fischer JE. The impending disappearance of the general surgeon. JAMA 2007;298(18): 2191-2193.
26. Stitzenberg KB, Sheldon GF. Progressive specialization within general surgery: adding to the complexity of workforce planning. J Am Coll Surg 2005;201(6): 925-932.
27. Bell PR. Vascular surgery: recognition at last. Vascular 2012;20(5): 241-242.
28. Soreide K. Trauma and the acute care surgery model--should it embrace or replace general surgery? Scand J Trauma Resusc Emerg Med 2009;17: 4.
29. Richardson JD. Training of general surgical residents: what model is appropriate? Am J Surg 2006;191(3): 296-300.
30. Debas HT. Surgery: a noble profession in a changing world. Ann Surg 2002;236(3): 263-269.

Appendices

Acknowledgements, Dankwoord
PhD portfolio
Curriculum Vitae auctoris

ACKNOWLEDGEMENTS, DANKWOORD

Opgelucht ben ik dat ik mijn promotie-onderzoek heb afgerond. Iedereen die op welke manier dan ook heeft bijgedragen aan dit proefschrift wil ik graag bedanken, een aantal personen in het bijzonder:

Geachte promotor, prof.dr. J.N.M. IJzermans, beste prof, u bent verantwoordelijk voor het initiëren van het promotietraject dat heeft geleid tot dit proefschrift. Hartelijk dank voor deze kans die u mij heeft geboden. Ik heb de afgelopen jaren enorm veel geleerd. De samenwerking verliep goed en uw visie en commentaar heb ik altijd als zeer waardevol ervaren. Uw korte e-mails met betrekking tot het afronden van dit proefschrift waren voor mij veelzeggend. Graag zou ik in de nabije toekomst nog eens met u aan tafel willen staan.

Geachte co-promotor, dr. P.T. den Hoed, beste Ted, na mijn oudste co-schap mocht ik in het Ikazia als arts-assistent aan de slag. Op de werkvloer werd mijn de liefde voor de chirurgie bevestigd. Je bent voor mij heel belangrijk geweest bij het maken van dit proefschrift en ik ben je zeer erkentelijk voor alles wat je in de afgelopen jaren voor me hebt gedaan.

Geachte leden van de leescommissie, prof.dr. M. de Hoog, prof.dr. I.H.M. Borel-Rinkes en prof.dr. L.P.S. Stassen, hartelijk dank voor het kritisch lezen en beoordelen van mijn proefschrift. Tevens wil ik prof.dr.ir. A.P.N. Themmen, prof.dr. J.L.C.M. van Saase, dr. P.K. Bos en dr. B.P.L. Wijnhoven danken voor hun bereidheid om als opponent zitting te nemen in de promotiecommissie.

Beste ROC-leden, (Ted) de Hoed, (Lijckle) van der Laan, (Erwin) van der Harst, (Maarten) van der Elst, (Guido) Mannaerts, (Imro) Dawson en (Bas) Wijnhoven, iedere maand mocht ik in een korte tijd mijn verhaal vertellen over de vorderingen van mijn proefschrift. Regelmatig had ik na afloop rode kaken en was ik niet altijd even gelukkig met het commentaar. Het heeft mij door de tijd heen moed gegeven en ben ik gesterkt om door te zetten en mijn bevindingen aan het papier toe te vertrouwen.

Beste Rosemarijn Steensma, hartelijk dank voor je hulp en support tijdens deze periode.

Iris Wallenburg. Dank voor je hulp bij het opstarten voor mijn onderzoek. Ondanks de vele uren, zijn we erin geslaagd het mixed-methods artikel te publiceren. Wat een verrassing dat het is bekroond met een prijs! Veel succes in je verdere carrière.

Beste co-auteurs, jullie bijdrage aan dit proefschrift spreekt voor zich. Hartelijk dank! In het bijzonder wil ik Reinier Timman bedanken voor zijn fantastische hulp bij de statistiek.

Beste secretaresses, dank voor jullie hulp bij de vele zaken die geregeld moesten worden gedurende het schrijven van dit proefschrift. Een speciaal woord van dank aan Carola Zandijk voor haar nauwe betrokkenheid bij de organisatie van deze promotieplechtigheid in de afgelopen maanden.

Chirurgen van het Amphia Ziekenhuis, 3 jaar geleden ben ik met de opleiding Heelkunde gestart in jullie ziekenhuis. Ik kan niet anders zeggen dan dat ik het enorm naar mijn zin heb en de verantwoordelijkheid en het vertrouwen dat jullie in mij hebben zeer waardeer. Ik ben er van overtuigd dat ik mij met behulp van jullie begeleiding de aankomende jaren zal ontwikkelen tot een goede (vaat)chirurg.

Chirurgen en (oud)arts-assistenten van het Ikazia Ziekenhuis, met veel plezier en enthousiasme heb ik anderhalf jaar als arts-assistent bij jullie gewerkt. Hartelijk dank voor deze leerzame tijd. Nog dagelijks pluk ik de vruchten van wat jullie mij tijdens deze periode hebben geleerd. Voor weekenddienst tijdens het jaarlijkse skiuitje houd ik mij graag aanbevolen.

Beste paranimfen,
Kevin, in de relatief korte tijd dat ik je ken ben je naast een fijne collega in de kliniek een goede vriend van mij geworden. Ik heb vele mooie herinneringen aan de cursussen en congressen die we de eerste jaren van de opleiding samen hebben bezocht, vooral de AO-cursus in Davos staat in mijn geheugen gegrift. Veel succes in het Erasmus MC!

Bo, beste medebewoner en buurman in Z-839. Dank voor het luisterend oor dat ik af en toe nodig had. Ik ben erg blij dat ik je heb leren kennen en dat jij op 5 juli naast me wilt staan. Jouw boekje is niet ver weg!

Mede-onderzoekers uit de Z-flat, veel dank voor alle hulp en gezellige praatjes. De spontane vrijdagmiddagborrels met liquid dinner zal ik nooit vergeten.

Beste vrienden, in het bijzonder Erik Westerink en Jan-Peter Boer. De “laatste loodjes” van mijn proefschrift heb ik de afgelopen jaren voltooid naast mijn reguliere werkzaamheden. Dit betekende dat we elkaar minder zagen dan voorheen. Nu dit proefschrift afgerond is, hoop ik meer tijd te hebben om elkaar te ontmoeten, gezellig samen te eten en op stap te gaan.

Familie, vrienden, collega’s en mensen die ik niet persoonlijk heb genoemd. Bedankt voor jullie interesse en meeleven.

Dear parents-in-law, Christian and Patricia, I am very grateful for your support and thorough reviews as a native speaker. I really appreciate the time and suggestions provided by e-mail and during your various visits to the Netherlands. We look forward to visiting New Zealand again but first we will meet in Japan.

Lieve papa en mama, lieve broertjes, heel veel dank voor jullie onvoorwaardelijke steun, het rotsvaste vertrouwen in mijn kunnen en een luisterend oor. Jullie zijn heel bijzonder voor me. Een betere familie dan jullie kan ik mij niet wensen. Ik hou van jullie!

Lieve Tania, het boek is klaar! Wat moet dit een enorme opluchting voor je zijn. Ik heb de afgelopen jaren ongelofelijk veel van je gevraagd. Je was geduldig, bood me ruimte, gaf steun en bent tevens verantwoordelijk voor het design en opmaak van dit proefschrift. Ik ben ongelooflijk trots op jou. We gaan samen met Lily-Rose nog heel veel genieten. Jullie zijn de allerliefste. I love you to the moon and back!

PHD PORTFOLIO

Name PhD student: Cornelis Johannes Hopmans
Erasmus MC Department: Department of Surgery
PhD period: July 2011 - June 2014

Promotor: prof.dr. J.N.M. IJzermans
Co-promotor: dr. P.T. den Hoed
Date defense thesis: 05-07-2017

1. PhD training

	Year	Workload (ECTS)
General courses		
- Implementatietraining Modernisering Medische Vervolgopleidingen (MMV)	2011	1.0
- Introduction to Clinical Research	2012	0.9
- Biostatistics for Clinicians	2012	1.0
- Research Integrity	2012	0.6
- Basic Introduction Course on SPSS	2012	1.0
Seminars and workshops		
- Modernisering Medische Vervolgopleidingen (MMV) "Werk in uitvoering"	2011	2.0
Oral presentations (0.5 points each)		
- European Society of Esophagology	2011	0.5
- Chirurgendagen	2012-2014	1.5
- European Society for Surgical Research (ESSR)	2013	0.5
- World Congress on Surgical Training (SURGICON)	2013	0.5
- International Conference on Surgical Education and Training (ICOSET)	2014	0.5
- Annual Clinical Congress of the American College of Surgeons	2014	0.5
Poster presentations (0.5 points each)		
- Chirurgendagen	2012,2014	1.0
- World Congress on Surgical Training (SURGICON)	2013	0.5
- International Association for Medical Education (AMEE)	2015	0.5

(Inter)national conferences (0.3 points/day)		
- Chirurgendagen	2012-2015	2.4
- European Society for Surgical Research (ESSR)	2013	1.2
- World Congress on Surgical Training (SURGICON)	2013	1.5
- International Conference on Surgical Education and Training (ICOSET)	2014	1.5
- Annual Clinical Congress of the American College of Surgeons	2014	1.5

2. Teaching

	Year	Workload (ECTS)
Lecturing		
- Lecturing at Department of Surgery in onderwijs- en opleidingsregio IV	2012-2013	3.5
- STZ opleidingssymposium OOR-ZWN, Delft	2013	2.5
- Vascular Rounds, Rotterdam	2014	2.0
Supervising practicals and excursions, Tutoring		
- Examinator Basic Life Support for medical students	2012-2014	1.5

CURRICULUM VITAE AUCTORIS

Cornelis Johannes (roepnaam: Niels) Hopmans werd geboren op 26 maart 1984 als zoon van Hans en Wil Hopmans. Hij groeide op in Tholen (Zeeland). In 2002 behaalde hij zijn gymnasiumdiploma aan het Moller Lyceum in Bergen op Zoom en in datzelfde jaar begon hij aan zijn studie Geneeskunde aan de Erasmus Universiteit Rotterdam.

Na het behalen van zijn artsexamen in 2009 werkte hij met veel plezier en enthousiasme anderhalf jaar als arts-assistent op de afdeling Heelkunde van het Ikazia Ziekenhuis te Rotterdam (dr. P.T. den Hoed). Tijdens deze periode werd zijn interesse in de Heelkunde bevestigd.

In zijn streven om chirurg te worden heeft hij zich vervolgens toegelegd op het verrichten van wetenschappelijk onderzoek aan de afdeling Heelkunde van het Erasmus Universitair Medisch Centrum Rotterdam (promotor: Prof.dr. J.N.M. IJzermans en co-promotor: dr. P.T. den Hoed). De resultaten van zijn onderzoek naar de recente veranderingen en ontwikkelingen binnen de opleiding Heelkunde zijn gebundeld in dit proefschrift.

Op 1 juli 2014 is hij gestart met zijn opleiding tot chirurg in Onderwijs- en Opleidingsregio Rotterdam (regio IV) onder supervisie van dr. L van der Laan (Amphia Ziekenhuis, Breda) en dr. B.P.L. Wijnhoven (Erasmus Universitair Medisch Centrum, Rotterdam).

In zijn vrije tijd voetbalt Niels met veel plezier samen met zijn 2 jongere broers Marien en Hans-Willem in het 1ste elftal van Tholense Boys. Niels woont samen met zijn vrouw Tania Hopmans - Pearse-Danker in Rotterdam. Samen hebben zij een prachtige dochter: Lily-Rose (10-7-2016).