Universidade do Minho
Escola de Medicina

Domenico Veneziano

Development of Endourology basic skills curriculum for urologists in training

junho de 2020

UMinho | 2020 Ana Rita Machado dos Santos **Development of Endourology basic skills curriculum for urologists in training**

Universidade do Minho
Escola de Medicina

Domenico Veneziano

Development of Endourology basic skills curriculum for urologists in training

Tese de Doutoramento
Doutoramento em Medicina

Trabalho efetuado sob a orientação do
Prof. Estevao Augusto Rodrigues de Lima

junho de 2020

Acknowledgements

To Professor Estevao Lima, Supervisor of this thesis, Innovator and Friend, for having encouraged me to embrace this research study, which changed my career and hopefully will impact thousands of Urology students from all around the world.

To Professor Juan Palou, Chairman of the European School of Urology, for having supported since the very beginning the renovation and widespread of the EAU hands-on training system with immense enthusiasm.

To ESU Training group, Ben Van Cleynenbreugel, Ali Goezen, Bhaskar Somani, for having provided full support and collaboration in the development and adoption of protocol and strategies.

Special Thanks to the EUREP hands-on training team for having helped in data collection, wide-spreading the EST s1 worldwide and having profoundly believed in this project.

To my family for their patience, love and comprehension, even in the toughest days of my career.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Development of endourology basic skills curriculum for urologists in training

Abstract

In the last decade, surgery training has radically changed with the implementation of practical training programs prior to surgery on Human. This training model has been accepted in the worldwide urological community and its effectiveness has been demonstrated in several studies. In Europe, the European Urological Association has developed and implemented a basic laparoscopic surgery skills curriculum which it has called EBLUS (European Basic Laparoscopic Urological Skills).

Given that in the area of endoscopic surgery there was no basic training curriculum yet, the aim of this thesis was to develop, validate and improve the first practical training protocol in Europe for Endoscopic Stone Treatment as basic skill in endourology.

As a result of this thesis, a training method for endoscopic stone treatment step 1 (EST s1) was developed to generalize it as a form of initial acquisition of endourology skills. Subsequently, this first training step was validated for its importance in the curriculum training of residents through collaborative work between various sections of the European Urological Association.

In order to improve surgical training, the Performance Improvement (PI) score was initially developed and validated during EBLUS sessions and during EST s1 sessions.

New curriculums and training methods are needed in endourological training prior to surgery on humans. Thus, this thesis can contribute to improving the endourology training program.

Keywords: Endourology, Hands-on Training, Simulation, Standardization, Training.

Desenvolvimento de um Curriculum de aptidões básicas em endourologia para os urologistas em formação

Resumo

Na última década, o treino em cirurgia mudou radicalmente com a implementação de programas de treino prático prévio à realização de cirurgia no Humano. Este modelo de treino tem vindo a ser aceite na comunidade urológica mundial e a sua eficácia tem sido demonstrada em vários estudos. Na europa, a Associação Europeia de Urologia desenvolveu e implementou um curriculum básico de aptidões em cirurgia laparoscópica que denominou de EBLUS (European Basic Laparoscopic Urological Skills).

Dado que na área da cirurgia endoscópica ainda não havia qualquer curriculum básico de treino, o objetivo desta tese foi desenvolver, validar e aprimorar o primeiro protocolo de treino prático na europa de tratamento endoscópico de cálculos como aptidão básica em endourologia.

Como resultado desta tese desenvolveu-se um método de 1ª etapa de treino no tratamento endoscópico de cálculos (EST s1) de forma generaliza-lo como forma de aquisição inicial de aptidões em endourologia. Posteriormente, esta 1ª etapa de treino foi validada a sua importância na formação curricular de residentes através de um trabalho colaborativo entre várias secções da associação europeia de urologia.

De forma a objetivar o melhoramento com o treino cirúrgico foi desenvolvido e validado o "Performance Improvement (PI) score" inicialmente durante as sessões de EBLUS e durante as sessões de EST s1.

Novos currículos e métodos de treino são necessários na formação endourologica antes da realização de cirurgia em Humanos. Esta tese pode contribuir assim para melhorar o programa de formação em endourologia.

Palavras-chave: EndoUrologia, Hands-on training, padronização, Simulação, treinamento.

Table of Contents

Abbreviations:

EAU: European Association of Urology

EST S1: Endoscopic Stone Treatment – Step 1 ESU: European School of Urology

ESUT: European Section of Uro-Technology EULIS: European Section of Uro Lithiasis

EUREP: European Urology Residents Education Programme HoT: Hands-on Training

Pi-score: Performance Improvement score

List of Figures

List of Tables

CHAPTER 1

Introduction

1. Introduction

1.1. History of surgical training and today's complex scenario

"Learning by doing" is a very effective methodology to acquire new skills. It's more engaging for the learner and offers the possibility to personally apply problem-solving strategies. As theorized by the American philosopher and pioneer in pedagogy John Dewey (1859 - 1952), learning should be "relevant and practical, not just passive and theoretical". Being curious about what's surrounding us is just one of the reasons why we love "doing". Replicating everyday situations is indeed one of the favorite activities of kids: cooking on a toy kitchen, driving little car replicas, dressing and acting as firefighters are just some of the hundred games that they love to play. These are the first steps of a long journey, that drives humans while they grow up and until the very last day of their lives. Simulating and playing is a great way not only to learn, but also to do it safely, with no harm for yourself or for the others. The first simulator in history, the Link Trainer (figure 1), was indeed designed to prevent errors in the most dangerous situation for U.S. Air Force pilots: landing on an aircraft carrier. It is still possible to watch an original footage on the internet [1] demonstrating how happy the officers were in the case of a mistake: that error that was showing up during a simulated session, was indeed avoiding a real accident and potentially saving the life of a pilot.

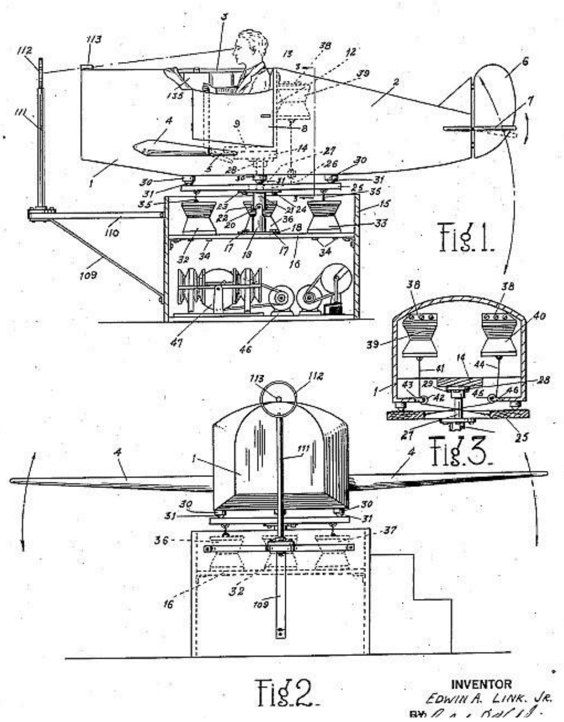


Figure 1: the link trainer. Technical drawing

Being able to make mistakes, understanding and applying problem-solving strategies with no harm for the patient, are some of the several critical needs of any surgeon in training. At the same time, simulation might be very useful also for those who simply want to stop and think, before approaching a difficult, unusual case.

Medical error, according to Makary and Daniel [2] is today the third leading cause of death in the US. Despite this “ranking” being criticized by many at the moment of publishing, deaths by medical errors have raised up from 90.000 per year in 1999, to 250.000 per year in 2016. The weight of these data is further increased by their cost: 20.8 billion US dollars due to complications in 2016, which raises up from 187 to 250 billions if we consider the loss in QALY (quality-adjusted life

years) for the registered deaths, with an average of 10 years of life lost per person [3]. Technical skills have a direct impact on clinical outcomes, with 2,5 times more readmissions, 3 times more complications and up to 5 times more deaths after surgeries completed by poor performers vs the top performers [4]. Despite the huge need for training, a study published in 2013 [5] revealed that participation in laparoscopy courses and fellowships is low and should be encouraged. The same data have been confirmed later on by a manuscript from 2016 [6], which analyzed Urology training in Italy. Above the 324 residents surveyed, very few completed a radical prostatectomy or its early steps, before the end of their residency. Even the data collected from the European Urology Residents Education Program (EUREP) are not depicting a brighter situation, as between the 1288 fifth-year residents sampled, just 32% were working in centers provided with surgical simulators. Even between those lucky ones (410 of 1288), just 17% declared to perform surgical training once a week, while others were dedicating less time to their personal surgical-skill growth. From these data, even if at first sight the problem might be related just to insufficient training facilities, on the other hand it seems to be related to a scarce interest of trainees into training itself. Indeed, not all training, not even hands-on training (HoT), is always engaging. This is the reason why in the last decade several improvements have been applied to make it more effective, appealing and accessible on a large scale. What is today called “EUREP HoT format” [7] depicts an easily replicable hands-on training set-up, that has been tested over time with repeatedly high-quality results. The European Association of Urology (EAU) started to officially deliver the first standardized HoT protocol in 2011 and, since then, over 40 countries have been

involved in the process, with the final aim of providing certified surgical education globally. This may lead on the long run to a complete change in the paradigm of medicine, as surgeons might start to

actually touch their first patient after having shown proper mastery of the skills required to do it, which is a critical step still missing. It could have a drawback on several aspects of medicine, not only related to patient satisfaction and safety, but also to costs, insurances and management. For all the aforementioned reasons, surgical training is a field that just started to show its potential and is destined to still grow a lot. Technology is progressively being integrated in different ways, impacting methodologies, assessment, visualization modalities and even psychological aspects. While many virtual simulators are already on the market, some start-ups have developed phone-apps [8] to experience full surgical procedures, in order to facilitate the understanding of procedural steps. At the same time, the most used training platforms are still organic, like lab animals and cadavers, while synthetic models are progressively being adopted. The possibility of "mixed reality integrations", meaning adding virtual layers over physical objects, could enrich the simulative experience.
In consideration of such a complex and relevant scenario, we wanted to focus this chapter on how HoT is evolving and what are its future applications, starting from the backstage of its early adoption and development. Let's discover how "learning by doing" is really applicable to Urology and how it can help us keeping our patients safer.

The modern era of surgical simulation started in 1990 with the development of a system built by NASA members Rosen and Delp, to practice tendon transplants

in gait disorders [9]. Surgical simulation can be defined as an exercise or device that enables the trainee to practice a surgical task several times and under safe test conditions. Minimal invasive surgery and laparoscopy started with the first steps in the late eighties and gained popularity at the beginning of the nineties. The technique has rapidly found indications in Urology, despite its application being initially debated because considered too challenging. Indeed, getting used to watching a bi-dimensional image on a screen, without any possibility to "feel" tissues and directly see them, was something totally new for the surgeon. Laparoscopy, with the special set of instruments required, was not easy and had a steep learning curve, which is still representing a challenge for the average urologist. These new and special technological breakthroughs were not really implantable in the classical surgical training programs and needed promotion of a change in surgical skills acquisition: new training programs and a new philosophy of teaching. The aim was designing attractive laparoscopic courses, with easy transferability to the clinical setting. A comprehensive training curriculum was missing in the first applied HoT programs. Courses were mostly organized and promoted by the early adopters with great variations from center to center. Rassweiler et al. were between the first to arrange training events on lab animals [10] and in particular on porcine model, that was allowing the closest scenario to actual surgery on humans due to the similar anatomic landmarks. Critics were by the time about going straight from animals to humans, so concepts like the LapTent [11] came out. This training device by Henkel et al. was consisting of a drape-folded rectangular metal frame, to be placed over the open wound of the patient (figure 2).

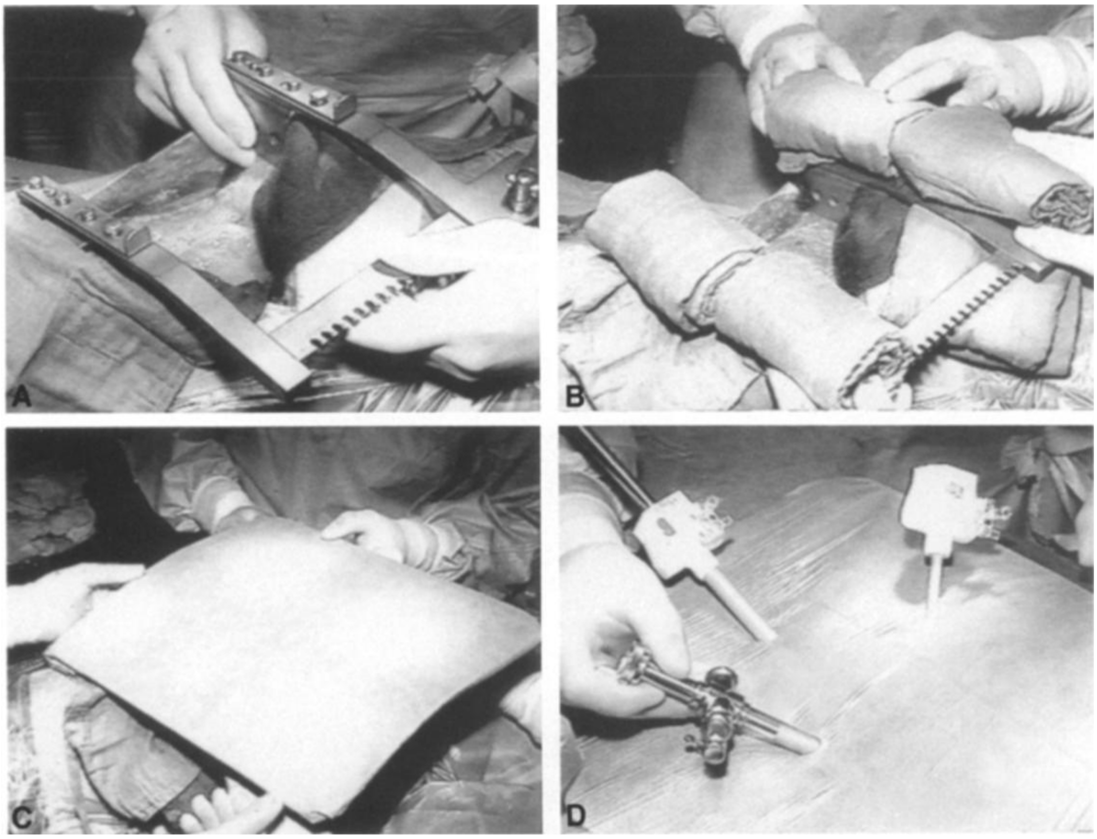


Figure 2: the LapTent by Henkel [11]

It allowed to take confidence with the mini-invasive approach in the attempt of "bridging the gap between open and laparoscopic surgery". Another common model used in the early 90s and still sometimes adopted, was a simple sponge. It was a clever way to teach how to apply intra-corporeal knots and avoided placing organic parts inside a lap trainer.

Similar initial modalities tended to disappear, while simulation-based training rapidly proved to be an effective method for acquiring laparoscopic skills. The patient was by the time still considered as part of the training, but with the aid of a mentor. In 1999 the EAU section of Uro-technology (ESUT) started to develop training programs and to organize HoTs for laparoscopic surgery. However, the qualification and certification of laparoscopic skills was still in an early phase and standardization was still low. The EUREP course by European School of Urology (ESU) started in 2003 and was delivered to last year urology residents. While in the beginning it was hosting just theoretical classes, HoTs were included since 2007. In the same years a 5-step training program for laparoscopic urologic skills was been adopted in the Netherlands. Out of the 5 tasks included in the curriculum, 4 were inspired by the American FLS (Fundamentals of Laparoscopic Surgery) protocol [12] and were applied with different timing and error assessment, in accordance with the European needs. Task 3, needle guidance, was developed by the Dutch project Training in Urology (TiU), which took care also of the validation of the entire protocol [13]. In 2011 this program was introduced at a European level as a pilot project for the residents during EUREP, under the name “European Basic Laparoscopic Urologic Skills”, or E-BLUS. EBLUS was initially composed by 5 tasks: peg transfer, circle cutting, needle transfer, knot tying and clipping vessel. After being delivered for the first time in 2011 in Prague during the EUREP course, in 2012 the protocol underwent some modifications: time cut-offs were slightly trimmed and task number 5 was deleted, due to the poor relevance of the task and to clips-related costs. In the same year the program was used to assess and certify the technical laparoscopic skills of the European residents. An optional task, Camera Handling [14], was introduced the same year for beginners who were mostly aiding as cameramen in their home institutions, to provide tips and tricks about the use of 30° optics. An online web based theoretical course, including history of minimal invasive surgery, laparoscopic instruments and safety, physiology and aspects of anesthesiology

and training, was added to the E-BLUS as a mandatory part to access the practical examination. This was intended to support HoT with a solid, standardized background knowledge. EBLUS progressively gained interest in Europe and in 2013 was officially performed at a national level for the first time in Italy, during the national convention of the Società Italiana di Urologia (SIU), in Riccione. A big improvement in the organization of the events has been the collection of feedbacks from participants, that showed continuous improvements since 2014 [7]. The program, proven to be easily replicable, is today considered as a model to be followed and has been delivered for training and assessment in over 40 countries worldwide (figure 3), reaching not only Europe, but also Asia, Africa and America.

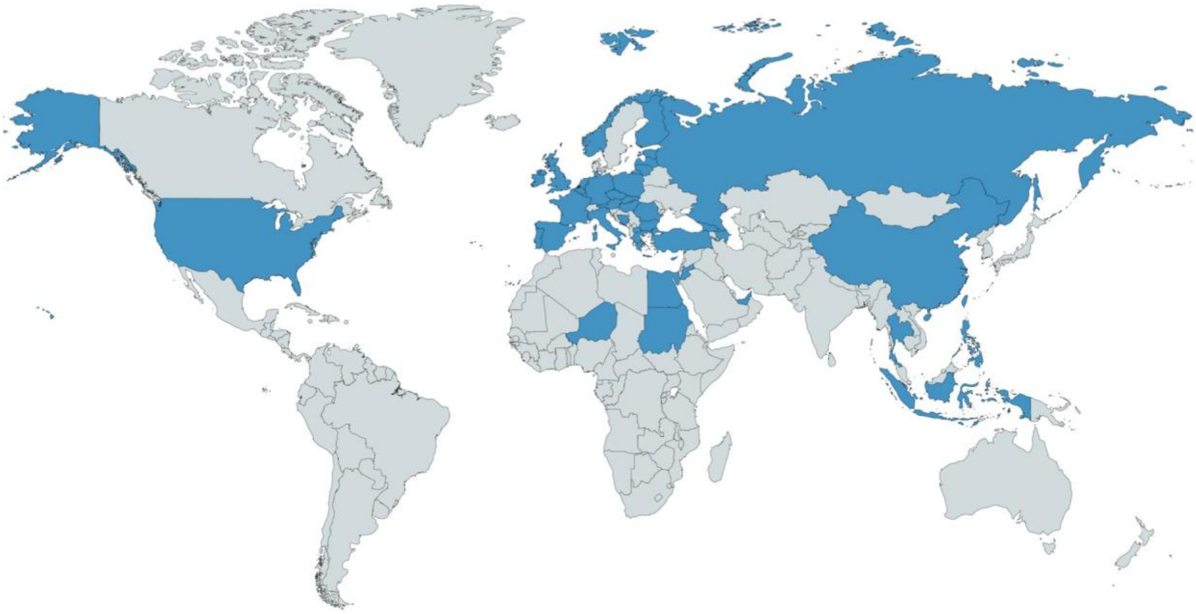


Figure 3: a map of the EBLUS widespread in 2018

The theoretical basic-knowledge part has been recently updated in accordance with literature and new explanatory videos have been included in the EAU official website.

Given the huge widespread of the EBLUS protocol and the high demand for stone treatment procedural training, the development of a dedicated standardized training protocol was needed, with the aim of guiding students and urologists in training with full respect of patient safety.

References

1) https://www.youtube.com/watch?v=Im_G2x4k6zQ
2) BMJ 2016; 353: i2139 doi: 1136/bmj. i2139 (Published 3 May 2016)
3) Andel, et al. in "The Economics of Health Care Quality and Medical Errors," Journal of Health Care Finance, Fall 2012
4) Birkmeyer, et al, NEJM, October 2013
5) BJU Int. 2013 Dec;112(8):1223-8. doi: 10.1111/bju.12410. Training of European urology residents in laparoscopy: results of a pan- European survey. Furriel FT1, Laguna MP, Figueiredo AJ, Nunes PT, Rassweiler JJ.
6) Eur Urol Focus. 2016 Jun 21. pii: S2405-4569(16)30066-9. doi: 10.1016/j.euf.2016.06.006. Urology Residency Training in Italy: Results of the First National Survey. Andrea Cocci et al.
7) Eur Urol Focus. 2018 Mar 14. pii: S2405-4569(18)30080-4. doi: 10.1016/j.euf.2018.03.002. The European Urology Residents Education Programme Hands-on Training Format: 4 Years of Hands-on Training Improvements from the European School of Urology.
8) Example: www.touchsurgery.com
9) Delp, S. L., Loan, J. P., Hoy, M. G., Zajac, F. E., Topp, E. L. andRosen, J. M.: An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. IEEE Trans Biomed Eng, 37:757, 1990
10) Rassweiler JJ, Henkel TO, Potempa DM, Frede T, Stock C, Günther M, Alken P. Laparoscopic training in urology. An essential principle of laparoscopic interventions in the retroperitoneum. Urologe A. 1993 Sep;32(5):393-402. German. PMID: 8212425
11) Henkel TO, Potempa DM, Rassweiler J, Manegold BC, Alken P. Lap simulator, animal studies, and the Laptent. Bridging the gap between open and laparoscopic surgery. Surg Endosc. 1993 Nov-Dec;7(6):539-43.PMID: 8273005
12) Fried GM. FLS assessment of competency using simulated laparoscopic tasks. Journal of Gastrointestinal Surgery. 2008;12(2):210-2.
13) Brinkman WM, Tjiam IM, Schout BMA, Muijtjens AMM, Van Cleynenbreugel B, Koldewijn EL, et al. Results of the European Basic Laparoscopic Urological Skills Examination. Eur Urol. 2014;65(2):490-6.
14) Veneziano D, Minervini A, Beatty J, Fornara P, Gozen A, Greco F, Langenhuijsen JF, Lunelli L, Overgaauw D, Rassweiler J, Rocco B, Sanchez Salas R, Shariat S, Sweet RM, Simone G, Springer

C, Tuccio A, Van Cleynenbreugel B, Weibl P, Cozzupoli P. Construct, content and face validity of the Camera Handling Trainer (CHT): a new E-blus training task for 30° laparoscope navigation skills World J Urol - DOI 10.1007/s00345-015-1657-6

1.2. Objectives of the study

The popularity and of laparoscopic hands-on training have greatly increased in the last decade, thanks to the demonstrated validity of the basic curriculum developed and widespread worldwide by the European Association of Urology. The aim of this project is developing and validating the counterpart of EBLUS (European Basic Laparoscopic Urological Skills) for the acquisition of basic endoscopic stone treatment skills. Thus, we planned this PhD project with the following points to achieve:

1) Understanding of the basic principles of endoscopic stone treatment and development of a novel curriculum
2) Analysis of the actual gold standard methodologies for teaching a hands-on training session
3) Adding content, construct and face validity evidence to the curriculum
4) Developing an objective assessment tool for performance improvement, which will be firstly applied to an already validated protocol as a benchmark
5) Applying the new tool the the novel curriculum in order to evaluate performance improvement of the participants.

CHAPTER 2

Experimental work

CHAPTER 2.1

Veneziano D., Ahmed K., Van Cleynenbreugel B., Gozen A., Palou J., Sarica K., Liatsikos E., Sanguedolce F., Honeck P., Alvarez-Maestro M.,

Papatsoris A., Greco F., Breda A., Somani B.

Development methodology of the novel Endoscopic stone treatment step 1 (EST s1) training/assessment curriculum

Endourol. 2017 Jul 10. doc: 10.1089/end.2017.0248

Development methodology of the novel Endoscopic stone treatment step 1 (EST s1) training/assessment curriculum: An international collaborative work by EAU sections

Domenico Veneziano[1], Kamran Ahmed[2], Ben Van Cleynenbreugel[3], Ali Gozen[4], Joan Palou[5], Kemal Sarica[6], Evangelos Liatsikos[7], Francesco Sanguedolce[5], Patrick Honeck[8], Mario Alvarez Maestro[9], Athanasios Papatsoris[10], Panagiotis Kallidonis[11], Francesco Greco[12], Alberto Breda[5], Bhaskar K Somani[13] and ESU training group.

1. Dept of Urology, University of Minho, Braga, Portugal
2. Dept of Urology, Guy's Hospital, London, United Kingdom
3. Department of Urology, University Hospital Gasthuisberg, Katholieke Universiteit Leuven, Louvain, Belgium
4. Department of Urology, Klinikum Heilbronn, Heilbronn, Germany
5. Department of Urology, Fundació Puigvert, Universitat Autònoma de Barcelona, Barcelona, Spain
6. Department of Urology, Dr. Lütfi Kǵrdar Kartal Research and Training Hospital, Istanbul, Turkey
7. Department of Urology, University of Patras, Patras, Greece
8. **Dept of Urology, Klinikverbund Südwest,** Sindelfingen, Germany
9. Dept of Urology, Hospital Universitario La Paz, Madrid, Spain
10. University Department of Urology, Sismanoglio General Hospital, Athens, Greece
11. Department of *Urology* University of Patras Rio, Patras Greece
12. Dept of Urology and mini-invasive surgery, Romolo Hospital, Crotone, Italy
13. Department of Urology, University Hospital Southampton NHS Foundation Trust, Southampton, UK;

Abstract

Background

Simulation based technical-skill assessment is a core topic of debate, especially in high-risk environments. After the introduction of the E-BLUS (European Basic Laparoscopic Urological Skills) exam for basic laparoscopy, no more technical training/assessment urological protocols have been developed in Europe.

Objective

We describe the methodology used in the development of the novel Endoscopic Stone Treatment step 1 (EST s1) assessment curriculum.

Materials and Methods

The “full life cycle curriculum development” template was followed for curriculum development. A Cognitive Task Analysis (CTA) was run to define the most important steps and details of RIRS, in accordance with EAU Urolithiasis guidelines.
Training tasks were created between April 2015 and September 2015. Tasks and metrics were further analyzed by a consensus meeting with the EULIS board in February 2016. A review, aimed to study available simulators and their accordance with task requirements, was subsequently run in London on March 2016. After initial feedback and further tests, content validity of this protocol was achieved during EUREP 2016.

Results

The EST s1 curriculum development, took 23 months. 72 participants tested the 5 preliminary tasks during EUREP 2015, with sessions of 45 minutes each. Likert-scale questionnaires were filled-out to score the quality of training. The protocol was modified accordingly and 25 participants tested the 4 tasks during the hands-on training sessions of the ESUT 2016 congress. 134 participants finally participated in the validation study in EUREP 2016. During the same event 10 experts confirmed content validity by filling-out a Likert-scale questionnaire.

Conclusion

We described a reliable and replicable methodology that can be followed to develop training/assessment protocols for surgical procedures. The expert consensus meetings, strict adherence to guidelines and updated literature search towards an Endourology curriculum allowed correct training and assessment protocol development. It is the first step towards standardized simulation training in Endourology with a potential for worldwide adoption.

Introduction

Simulation based technical-skill assessment has been a core topic of debate, especially in high-risk environments. It was introduced in the aviation field in the early 30s with the "Link Trainer" [1]. Simulation training allows for safe and methodological acquisition of skills necessary for trainees, thereby also shortening their learning curve in this process. The first urology-specific curriculum came out in 2011, when the basic laparoscopic skills (E-BLUS) [2] exam was delivered for the first time by the European Association of Urology (EAU) and subsequently adopted worldwide. After the introduction of the E-BLUS exam, which was heavily inspired by the fundamentals of laparoscopic surgery (FLS) study carried-out by the American Society of Gastro-Enterological Surgeons (SAGES) [3], no more technical training/assessment urological protocols have been developed in Europe. With a recent surge in Endourology and stone treatment, a development of dedicated protocol-based training and assessment relating to endourological techniques was essential. While an OSATS-based (Objective Structured Assessment of Technical Skills) assessment tool for Ureteroscopy (URS) skills has been already described [4], no structured training curriculum has yet been developed.

Objective

Aim of this paper is to describe the methodology we followed to develop the EST s1, the first step of a novel modular training/assessment curriculum for Endoscopic Stone Treatment (EST). Our goals were to develop a set of replicable, standardized, low-cost exercises, providing objective assessment and applicability to 45-60 minutes hands-on training sessions. Given the structured pathway followed, our process could work as an example for developing new training protocols with highly reliable methodology and evidence.

Materials and Methods

The "full life cycle curriculum development" template (table 1), described by Richard Satava, was followed for curriculum development. As described by the author [5], the process starts by defining the outcomes and metrics; and ends with the certification that the outcomes, as planned at the beginning, have been properly achieved. The name "full life cycle" derives straight from this concept. Our whole development process started in November 2014 and is today at the stage of validation.

WHAT	**Outcomes & Metrics**	**Curriculum development**	**Simulator development**	**Validation studies**	**Implement Survey Training Certification**	**Issue Certification**
HOW	**Consensus Conference**	**Standard curriculum template**	**Engineering physical simulator**	**Standard validation template**	**Current procedures**	**Issue Mandates and Certificates**
EST s1	**nov 2014 YAUWP CTA feb 2015 ESUT meeting Davos**	**apr 2015 cognitive/HoT development sept 2015 preliminary test EUREP15 feb 2016 EULIS consensus**	**mar 2016 simulator review jun 2016 simulator adaptation jul 2016 simulator test ESUT16**	**sept 2016 validation EUREP16 Dec 2016 EULIS consensus**	**mar 2017 final tests EAU17 2017 ART in flexible**	**sept 2017 first exams EUREP15**

Table 1: full life cycle curriculum development.

Phase 1: Outcomes and Metrics

Prior to defining outcomes and metrics as described by the template, a Cognitive Task Analysis (CTA) [6,7] was run by the Endourology & Stone Treatment group of the Young Academic Urologist Working Party (YAUWP).

The CTA, as described by Clark and Estes, aims at defining the most important steps and details of a procedure and is the most appropriate way to objectively analyse it. In this case its goal was to study the details of Retrograde Intra-renal Surgery (RIRS), so CTA was run in accordance with EAU Urolithiasis guidelines [8] and in parallel with a focused updated literature review. Pre-operative, procedural, continuous variables and completion details were analysed in relation to each phase of the technique: cystoscopy, ureteroscopy, and stone fragmentation. Furthermore, indications, contraindications, equipment, patient positioning, procedural steps and “do’s and don’ts” were described in relation to each of the aforementioned procedural phases.

The CTA led to a complete description of the procedure, as suggested by the EAU guidelines, and defined its main steps, as depicted in table 2. Given the detailed procedural steps, we defined a modular training system, as described in fig. 1: Basic skills with focus on navigation and basic use of the operative channels, Intermediate skills with focus on lithotripsy techniques and Advanced skills with focus on complete endoscopic lithotripsy procedures. This protocol underwent a first consensus by Delphi method with ESUT experts in February 2015.

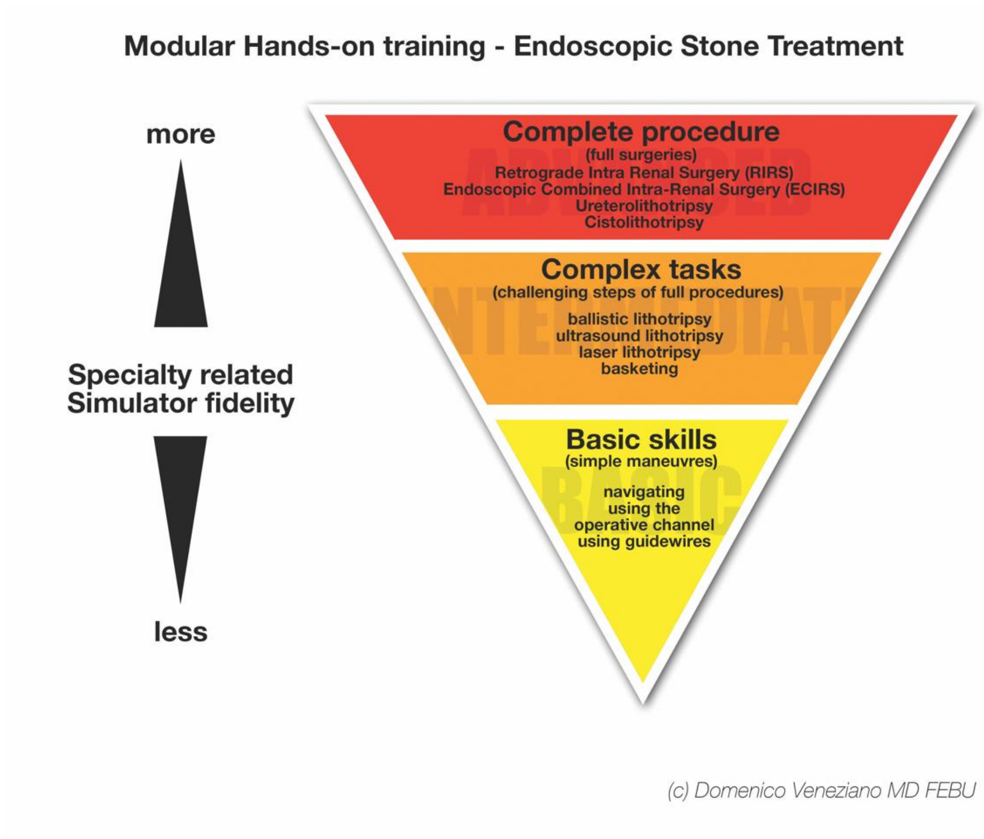


Figure 4: Modular hands-on training template for Endoscopic Stone Treatment.

Retrograde Intra-Renal Surgery procedural steps.	
Pre-operatory phase *1. Check patient-related details (correct patient, indication, allergies, culture, antibiotics, anticoagulants)* *2. Check material-related details (equipment up to date, presence materials, settings diathermy, irrigation fluid)* *3. Patient preparation (positioning, disinfection, sterile drapes)*	**Continuous variables procedural phase** **1. Change of instruments during procedure** **2. Orientation in the bladder** **3. Regulation of irrigation and emptying bladder** 4. Maintenance of visibility 5. Intra-renal pressure awareness 6. Identification of urotelial injuries
Procedural phase (RIRS) **1. Assembling instruments and connecting tubes** **2. Adjustment of light settings, focus camera, white balance** **3. Instillation of lubricant into meatus and introduction of the cystoscope** **4. Inspection bladder including orientation, identification of orifices and eventual bladder tumors** **5. Insertion of ureteric catheter** **6. Retrograde pyelography** **7. Ureteral mapping (identification of stones/strictures/filling defects)**	**Completion phase** *1. Documentation of stone-free status/residual fragments* **2. Removal of instruments** *3. Ensuring bladder emptying* *5. Debriefing (check count materials and stone-specimen/biopsy, discussion complications and postoperative policy)* *6. Registration (operating report, eventual pathology file, patient file, financial registration)*

8. X-ray guided placement of the guidewire (safety guidewire, X-ray safety precautions) **9. Semi-rigid ureteroscopy (inspection of the ureter)** 10a. In case of a ureteric stone, proceed with fragmentation (laser/ballistics) **10b. In case of a renal stone, placement of the working guidewire (through the semi-rigid ureteroscope)** **11. Placement of the access sheath (choosing the optimal size and length)** **12. Insertion of the flexible ureteroscope** **13. Inspection of calices** 14a. Insertion of the laser fiber (laser safety precautions, e.g. Glasses) 14b. Introduction of biopsy forceps in case of suspected lesions 15. Stone fragmentation/dusting, and basketing as appropriate 16. Double check under fluoroscopy and endoscopy for residual fragments 17. Stent placement if clinically indicated	

Table 2: RIRS procedural steps

Phase 2: Curriculum development

This phase was carried out in strict accordance with the CTA and was aimed to help the development of step 1 of the curriculum (basic skills). Training tasks were created from April 2015, until September 2015. "Indication and contraindication" information were used to structure the theoretical part of the trials, "equipment" was useful to exactly define the correct tools for each exercise, "patient positioning" and "procedure steps" were noted to define simulator requirements, while "do's and don'ts" were critical to define errors and metrics for each training task. The first set of exercises was composed as follows: task 1, flexible cystoscopy; task 2, safety guidewire placement with rigid cystoscope; task 3, rigid ureteroscopy and working guidewire placement; task 4, access sheath placement; task 5, flexible ureterorenoscopy. After a preliminary definition of the tasks, these underwent the first test during "European Urology Residents Education Programme (EUREP)" 2015. The aims of this test were to collect tutor feedback about the possibility of completing the defined protocol during a standard 45-minutes hands-on training session. The preliminary hands-on training step-1 protocol was delivered 72 times on 4 stations, with 4 expert tutors. No major issue was reported during the preliminary test. Likert scale-based quality feedback questionnaires were collected from the participants, with focus on several aspects of this training session (fig. 2): duration of the session, new skills acquired, basic skills improvement, expectations and overall evaluation of the course. Scores went from 1 (poor) to 5 (very good).

The results were collected along with a detailed description and refinement of these tasks. The tasks and metrics were further analyzed by a consensus meeting with the “EAU Section of Urolithiasis (EULIS)” board in February 2016. Following the feedback given by experts, these were reconfigured into a total of 4 tasks (table 3): task 1, flexible cystoscopy; task 2, rigid cystoscopy; task 3, semi-rigid ureteroscopy; task 4, flexible ureteroscopy.

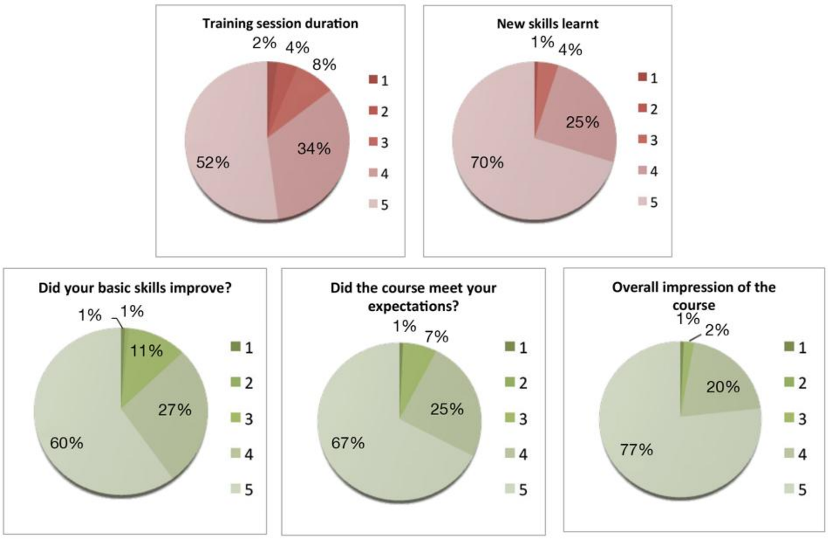


Figure 5: quality feedback questionnaires

Phase 3: Simulator development

Based on the simulator requirements of the final task-set, a review of the existing simulators was run at Guy’s Hospital, London, on March 2016. This was done in full overview of the information collected with the CTA at the beginning of the whole process. The simulators available for the trial were: Uro-Mentor (Simbionix, fig 3.A), Endo-Uro Trainer (Samed, fig 3.B), Advanced Scope Trainer (Mediskills, fig 3.C), Uro-Scopic Trainer (Limbs and Things, fig 3.D), K-Box (Coloplast, fig 3.E), KUB Model (University of Minnesota, fig 3.F). Once all the simulators had been tested, each experienced panel member independently filled out a simulator evaluation questionnaire. Panellists were asked to score with a 5-point Likert scale quality of the simulators in relation to different manoeuvres and characteristics: anatomic resemblance, bladder Visualization, instrument handling, ureteral catheterisation, ureteric navigation, semi-rigid URS and flexible URS. They were also given space for free-text comments. A final round-table discussion was then conducted to summarize their collective thoughts and findings.

Based on the results collected, two simulators (Endo-Uro Trainer by SAMED and K-Box by Coloplast) were chosen as the best fitting the needs of the protocol. The chosen models underwent some slight adaptation to allow full adoption of the task rules. The silicon bladder of the Endo-Uro Trainer was added with black marker dots to be used as targets. The K-box was considered in one standardized configuration, to allow for rotation, in-out and up-down movements with the flexible ureteroscope.

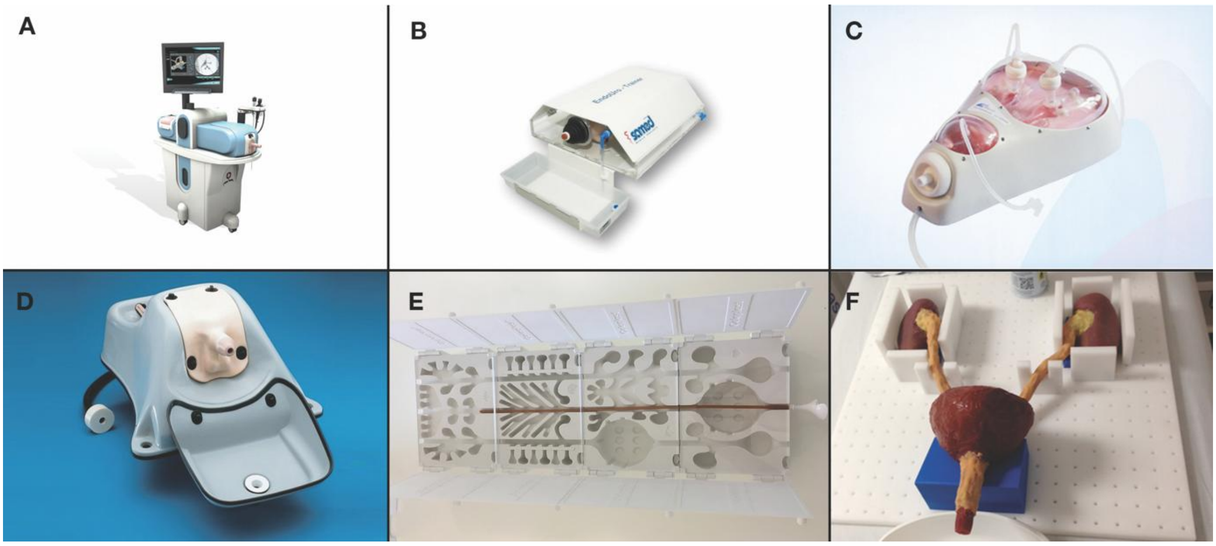


Figure 6: simulators tested

Phase 4: Validation

The custom-modified simulators were used for the first official test hands-on training sessions in Athens during "EAU Section of Uro-Technology (ESUT)" congress 2016, where they were used for a total of 25 one-hour sessions. Feedback from tutors was collected about the reliability of the simulators. Following the suggestions given during the tests, the black marker dots were replaced with 3mm black marker balls on the wall of the bladder models (fig 4), while K-boxes were provided with numbered cavities (fig 5), which allowed for an even more standardized task. After the described development process, preliminary validation of the protocol was run during EUREP 2016. In order to assess content validity of the developed protocol, ten experts from high volume stone treatment centres were asked to fill-out a Likert-score questionnaire, focused on each single task and its characteristics.

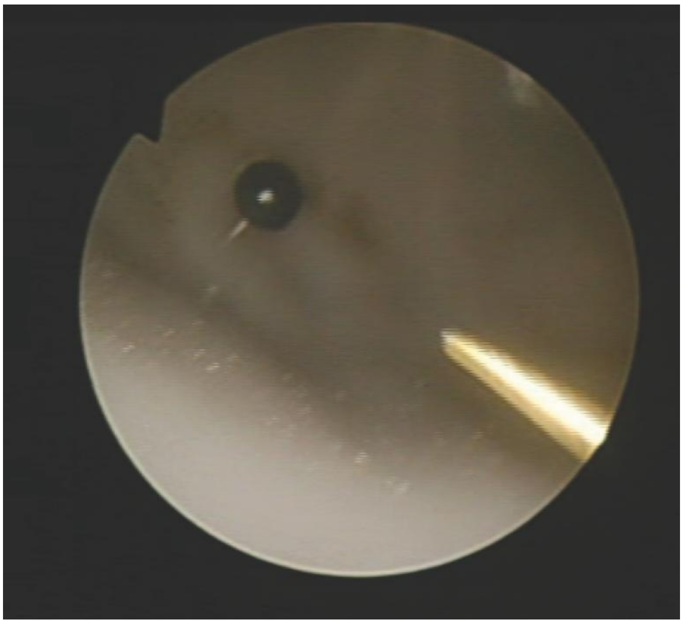


Figure 7: 3mm black marker balls on the wall of the bladder models

Figure 8: numbered cavities on kbox.

Results

The EST s1 curriculum development, from early data collection to validation, took a total of 23 months and involved EAU sections (ESUT, EULIS) and one YAUWP group, with guidance and coordination from the European School of Urology (ESU)/ESUT training group. The CTA, a detailed word document (6165 words) was structured in 5 paragraphs dedicated to the different parts of the Retrograde Intra-Renal Surgery (RIRS) procedure. Based on the CTA, a list of 3 pre-operative steps, 17 procedural steps and 5 completion steps was developed, defining the critical procedural phases to be simulated. This list was used to divide basic skills from intermediate tasks and cognitive contents (table 2). Seventy-two participants tested the 5 preliminary tasks during EUREP 2015, with sessions of 45 minutes each and the tutors reported no procedural issues. 86% of the trainees scored the training session duration as "good" or "very good". 95% of them stated that they acquired new skills during the session (fig. 2), while 87% declared that their basic skills improved during the course. The course, based on the preliminary EST s1 curriculum, met the expectations of 92% of the participants, who scored it as overall "good" in 20% of cases and as "very good" in 77% of cases. The 5 tasks curriculum was further optimized and after feedback it was readjusted to 4 tasks. During the simulator review phase, EndoUro Trainer by SAMED and the KUB model by University of Minnesota collected the highest scores based on the set simulator requirements (table 4). Due to the nature of these tasks, K-Box was selected just for task 4 (flexible ureteroscopy).

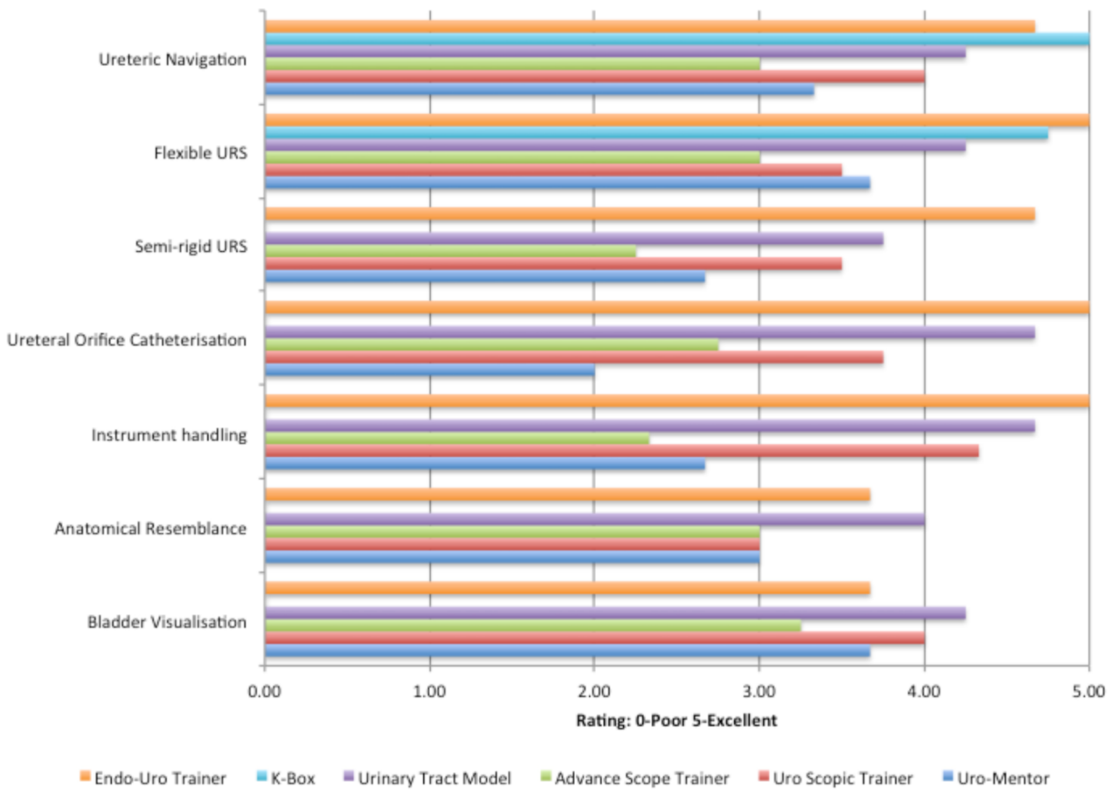


Table 3: results from simulator test

Subsequently, the new EST s1 protocol with 4 tasks were tested by 25 participants during the hands-on training sessions of the ESUT 2016 congress, and 134 participants with different expertise ranging from 0 to >1000 real-life URS cases finally participated in the validation study in EUREP 2016. The 10 experts involved (average of 23.7 URS stone treatment procedures/month) scored individually each task from the final task list (table 5). The statement "EST s1 should be used for basic URS training & assessment" was scored with a mode of 5 (±0,3), confirming the quality of the development process. Experts suggested that the curriculum should to be applied to the 3rd year of residency (±0,8). All tasks were scored as "valid basic training tools" with a mode of 5 (task1 ±0,6; task 3 ±0,5; task 4 ±0,4) excepted for task 2, scored with a mode of 4 (±0,5). The experts considered all tasks as proper parts of the curriculum (mode 5).

Question	Mode (±avg SD)
Task 1 - flexible cystoscopy	
It is a valid basic training tool	5 (±0,6)
It correctly represents a flex cystoscopy	4 (±0,5)
It should be part of the EST s1	5 (±0,5)
Task 2 - rigid cystoscopy	
It is a valid basic training tool	4 (±0,5)
It correctly represents a rigid cystoscopy	5 (±0,4)
It should be part of the EST s1	5 (±0,5)
Task 3 - semi-rigid URS	
It is a valid basic training tool	5 (±0,5)
It correctly represents a semi- rigid URS	4 (±0,4)
It should be part of the EST s1	5 (±0,5)
Task 4 - flexible URS	
It is a valid basic training tool	5 (±0,4)
It correctly represents a flex URS	5 (±0,6)
It should be part of the EST s1	5 (±0,5)
Overall rating	
It should be used for basic URS training & assessment	5 (±0,3)
When would you suggest the EST s1 training/exam to be performed?	
Year of residency	3 (±0,8)

Table 4: expert scores

Discussion

Meaning of our study

The new EST s1 protocol is the first such structured and streamlined Endourology curriculum simulation protocol, which has had collaborative input from various EAU sub-sections. The methodology has been done as a step-by-step replicable process following CTA, curriculum building, simulator review/development and feedback based protocol refinement. The result is a successful full cycle Endourology curriculum development, which has now undergone formal validation process.

Strengths and Weakness of our study

The strength of our study is the use of recognised methodology with a phased curriculum development. The CTA gave a solid background to the whole process, with a strong connection to the best achievable standards and guidelines. This allowed for a detailed curriculum building, which gave the training tasks an intrinsic validity, even before the popular "validation process", as depicted by several authors [9,10]. Indeed, even the preliminary task set tested during EUREP 2015 achieved high quality scores from the participants involved. Moreover, the process defined brought to the identification of specific simulator requirements, which lead to testing and scoring various simulators already available on the market and to finally modifying them in order to meet the requirements of the training curriculum.

This whole process successfully allowed the clinicians to wisely choose what was best fitting their needs, instead of simply using the plethora of available simulators. We consider this as one of the most important goals of our work, which should be considered as a priority for any curriculum development process.

In consideration of the expert involvement and of the scientific background given by the CTA, we acknowledge that our basic protocol has already achieved content validity, which has been anyway confirmed by the scores given by experts during EUREP 2016. Our statement is strengthened by the concept of "basic" surgical training curriculum, which is a set of basic skills that can be easily found or replicated in any procedure of similar nature. The analogy is equivalent to the E-BLUS, which contains simple skills such as cutting, bi-manual dexterity, knot tying, that are common to any laparoscopic procedure in urology, general Surgery and gynaecology. The same way, EST s1 tasks focus on navigation and basic use of operative channels, skills that can be found in urological endoscopy, but also in gastrointestinal or gynaecological endoscopy. Moving to more complex procedural steps, training becomes more specialty specific, just like happens for laparoscopy.

For complete endourology training, technical skills with a validated curriculum need to be supplemented with theoretical background, patient specific information and non-technical skills [7], which can help in training and assessment of communication skills in high stress or emergency situations. Cognitive information about the procedure were analysed and collected in a dedicated theory module by the educational group of the European Section of Urolithiasis, following the latest guidelines on this topic. The module will be integrated inside the curriculum as an addendum to the technical part. Non-technical skills related to the EST s1 protocol have not yet been analysed, even though their assessment can improve behaviour, team working and communication in these anxiety driven stressful situations [11,12]. Up to now, no weaknesses have been identified in the process, which appeared to flow smoothly and in full collaboration between different entities. The data collection from EUREP16 will eventually confirm the quality of the whole study, by integrating the already achieved content validity with face and construct validity.

Area of Future research

Although EST s1 is the first step in the integration of simulation in the Endourology curriculum, validation and further work with intermediate and advanced steps will be necessary for a comprehensive curriculum. Once this is established, subsequent work should involve fellowship in a recognised training programme in an Endourology unit.

From the information collected with the CTA and thanks to the preliminary expert consensus meetings, we already know that the intermediate step will focus on stone treatment, while advanced step will focus on full procedures and complication mastery. This work and the acquired methodology will allow for a faster development of the training protocols. Simulators enhance acquisition of skills thereby improving surgical training. The optimal duration and level of training needs to be targeted on trainee requirements and available resources. A modular training structure using low and high fidelity simulators that is realistic yet cost effective seems to be the best model for increased uptake and worldwide adoption.

Conclusion

The process summarized in this paper is a reliable and replicable methodology that can be followed to develop training/assessment protocols for any surgical procedure. The expert consensus meetings, strict adherence to guidelines and updated literature search towards an Endourology curriculum have allowed successful the achievement of content validity for the EST s1 training and assessment protocol. It is the first step in standardized simulation training in Endourology with a potential for worldwide adoption.

References

1) Link Trainer, *Wikipedia*: https://en.wikipedia.org/wiki/Link_Trainer (accessed 30 July 2016)

2) Brinkman WM, Tjiam IM, Schout BMA, Muijtjens AMM, Van Cleynenbreugel B, Koldewijn EL, et al. Results of the European Basic Laparoscopic Urological Skills Examination. *Eur Urol.* 2014;65(2):490-6.

3) Fried GM. FLS assessment of competency using simulated laparoscopic tasks. *Journal of Gastrointestinal Surgery.* 2008;12(2):210-2.

4) Argun OB, Chrouser K, Chauhan S, et al. Multi-Institutional Validation of an OSATS for the Assessment of Cystoscopic and Ureteroscopic Skills. J Urol. 2015 Oct;194(4):1098-105.

5) Satava RM, Gallagher AG. Next Generation of Procedural Skills Curriculum Development: Proficiency-based progression. JHS, 2015, Vol 3, Issue 4, Page 198-205

6) Clark R, Estes F (1996) Cognitive task analysis. Int J Educ Res 25:403–417

7) Hananel D, Sweet RM. Simulator Development – from Idea to Prototype to Product. In: R. Aggarwal JKJC-B, editor. *ACS Principles and Practice for Simulation and Surgical Education Research*. Chicago, IL: American College of Surgeons; 2015.

8) Türk C, Neisius A, Petrik A, Seitz C, Skolarikos A, Tepeler A, Thomas K
Guidelines Associates: Dabestani S, Drake T, Grivas N, Ruhayel Y. 2014 publication of the EAU Urolithiasis Guidelines.

9) McDougall EM. Validation of Surgical simulators. J Endourol 2007 Mar, 21(3): 244-7

10) Fried GM, Feldman LS, Vassiliou MC, Fraser SA, Stanbridge D, Ghitulescu G, Andrew CG. Proving the value of simulation in laparoscopic surgery. Ann Surg. 2004 Sep; 240(3):518-25; discussion 525-8.

11) Mishra A, Catchpole K, McCulloch P. The Oxford NOTECHS System: Reliability and validity of a tool for measuring teamwork behaviour in the operating theatre [Internet]. Qual. Saf. Health Care 2009, 18:104–8

12) Wood TC, Raison N, Haldar S, et al.: Training Tools for Nontechnical Skills for Surgeons—A Systematic Review. J. Surg. Educ. 2016, doi:10.1016/j.jsurg.2016.11.017.

CHAPTER 2.2

Somani B K, Van Cleynenbreugel B, Gozen A, Palou J, Barmoshe S, Biyani S, Gaya JM, Hellawell G, Pini G, Rodriguez F, Sanchez Salas R, Macek P, Skolarikos A, Wagner C, Eret V, Haensel S, Siena G, Schmidt M, Klitsch M, Vesely S, Ploumidis A, Proietti S, Kamphuis G, Geraghty R, Veneziano D.

The European Urology Residents Education Programme Hands-on Training Format: 4 years of Hands-on Training Improvements from the European School of Urology.

Eur Urol Focus. 2018 March 14 doi: 10.1016/j.euf.2018.03.002

The European Urology Residents Education Programme Hands-on Training Format: 4 Years of Hands-on Training Improvements from the European School of Urology

Bhaskar K. Somani [a,*], Ben Van Cleynenbreugel [b], Ali Gozen [c], Jaun Palou [d], Sas Barmoshe [e], Shekhar Biyani [f], Josep M. Gaya [d], Giles Hellawell [g], Gio Pini [h], Faba R. Oscar [d], Rafael Sanchez Salas [i], Petr Macek [j], Andreas Skolarikos [k], Christian Wagner [l], Viktor Eret [m], Stephen Haensel [n], Giampaolo Siena [o], Marek Schmidt [p], Max Klitsch [q], Stepan Vesely [r], Achilles Ploumidis [s], Silvia Proietti [h], Guido Kamphuis [t], Theodore Tokas [u], Rob Geraghty [a], Domenico Veneziano [v,w,x]

Department of Urology, University of Southampton, Southampton, UK;

Department of Urology, University Hospital Gasthuisberg, Katholieke Universiteit Leuven, Leuven, Belgium;

Department of Urology, SLK Klinikum Heilbronn, Heilbronn, Germany;

Department of Urology, Fundaciò Puigvert, Universitat Autònoma de Barcelona, Barcelona, Spain;

Department of Urology, Erasme Hospital, University Clinics of Brussels, Brussels, Belgium;

Department of Urology, St. James's University Hospital Leeds Teaching Hospitals NHS, Leeds, UK;

Department of Urology, The North West London Hospitals NHS Trust, Imperial College Healthcare NHS, London, UK;

Department of Urology, Ospedale San Raffaele–Turro, Milan, Italy;

Department of Urology, Institut Mutualiste Montsouris, Paris, France;

Department of Urology, Charles University, Prague, Czech Republic;

Department of Urology, University of Athens, Athens, Greece;

Department of Urology and Oncology, St. Antonius Hospital, Gronau, Germany;

Department of Urology, University Hospital Pilsen, Pilsen, Czech Republic;

Department of Urology, Havenziekenhuis, Rotterdam, The Netherlands;

Department of Urology, Azienda Ospedaliero Universitaria di Careggi, Florence, Italy;

Department of Urology, University Hospital Motol, Prague, Czech Republic;

Department of Urology, St. John of God Hospital, Vienna, Austria;

Department of Urology, Charles University 2nd Faculty of Medicine and University Hospital Motol, Prague, Czech Republic;
Department of Urology, Athens Medical Centre, Athens, Greece;
AMC University Hospital, Amsterdam, The Netherlands;
Department of Urology and Andrology, General Hospital, Hall in Tirol, Austria;
Department of Urology and Kidney Transplant, Grande Ospedale Metropolitano, Reggio Calabria, Italy;
School of Medicine, Hofstra University, New York, NY, USA;
Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal

Article info

Article history:
Accepted March 5, 2018

Associate Editor: Christian Gratzke

Keywords: Training Laparoscopy Endourology Ureteroscopy Simulation

Abstract

Background:

The European School of Urology (ESU) started the European Urology Residents Education Programme (EUREP) in 2003 for final year urology residents, with hands-on training (HOT) added later in 2007.

Objective:

To assess the geographical reach of EUREP, trainee demographics, and individual quality feedback in relation to annual methodology improvements in HOT.

Design, setting, and participants:

From September 2014 to October 2017 (four EUREP courses) several new features have been applied to the HOT format of the EUREP course: 1:1 training sessions (2015), fixed 60-min time slots (2016), and standardised teaching methodology (2017). The resulting EUREP HOT format was verified by collecting and prospectively analysing the following data: total number of participants attending differ- ent HOT courses; participants' age; country of origin; and feedback obtained annually.

Results and limitations:

A total of 796 participants from 54 countries participated in 1450 HOT sessions over the last 4 yr. This included 294 (20%) ureteroscopy (URS) sessions, 237 (16.5%) transurethral resection (TUR) sessions, 840 (58%) basic laparoscopic sessions, and 79 (5.5%) intermediate laparoscopic sessions.

While 712 residents (89%) were from Europe, 84 (11%) were from non-European nations. Of the European residents, most came from Italy (16%), Germany (15%), Spain (15%), and Romania (8%). Feedback for the basic laparoscopic session showed a constant improvement in scores over the last 4 yr, with the highest scores achieved last year. This included feedback on improvements in tutor rating ($p = 0.017$), organisation ($p < 0.001$), and personal experience with EUREP ($p < 0.001$). Limitations lie in the difficulties associated with the use of an advanced training curriculum with wet laboratory or cadaveric courses in this format, although these could be performed in other training centres in conjunction with EUREP.

Conclusions:

The EUREP trainee demographics show that the purpose of the course is being achieved, with excellent feedback reported. While European trainees dominate the demographics, participation from a number of non-European countries suggests continued ESU collaboration with other national societies and wider dissemination of simulation training worldwide.

Patient summary:

In this paper we look at methodological improvements and feedback for the European Urology Residents Education Programmehands-on-trainingoverthelast 4 yr.

Introduction

Minimally invasive surgery (MIS) started its true journey in the late 1980s and has seen an exponential rise in the last two decades [1,2]. With the evolution of minimisation and technological advances, there has been great variation in surgical training standards, leading to a need for training, simulation, and a structured curriculum. Endourology hands-on training (HOT) in a simulated environment started in the late 1990s but lacked a comprehensive training curriculum [3,4]. This led to the establishment of training protocols to structure and streamline training requirements and delivery [4–6]. Simulation training has gained momentum and there has been a huge rise in the number of simulators available, with trends suggesting an increase in the number of papers in this area published over the last two decades [7]. The European School of Urology (ESU) started the European Urology Residents Education Programme (EUREP) in 2003 for final year urology residents, with HOT added later in 2007. The motto in 2007 was "Sharpen your skills at the Dry-Lab courses which offer hands-on-interaction with state-of-the-art- equipment" and the trainee/tutor ratio was 3:1. Over the next few years, online course material was made available before the course for trainee preparation. This included history of endourology,

instruments, physiology, aspects of anaesthesiology, safety, and training. The first standardisation of the training methodology came in 2011 with a pilot European Training in Basic Laparoscopic Urological Skills (E-BLUS) examination [5]. At the time, the course provided 15 training stations for laparoscopy, with a trainee/tutor ratio of 2:1. Teaching sessions varied from 80 to 120 min. In 2014 quality feedback questionnaires were introduced, together with an additional basic laparoscopy task and a camera handling trainer, which was made available on each laparoscopy station. In the following years more effort was put in to standardise and create a training model that could be easily reproduced
outside the course in Europe and internationally.

We wanted to understand whether modifying the structure of HOT sessions regarding the duration of these sessions, the number of participants, the trainee/trainer ratio, and the teaching methodology resulted in a change in quality percep- tion by the trainees. In this study we analysed data collected from EUREP to provide individual quality feedback, trainee demographics, and the geographical reach of EUREP in relation to the methodological improvements adopted over the last 4 yr.

Materials and methods

Data for EUREP were prospectively analysed over the 4 yr from Septem- ber 2014 to October 2017. Several novelties have been applied to the HOT format of the EUREP course, including 1:1 training sessions (from 2015), fixed 60-min time slots (from 2016), and a standardised teaching meth- odology (from 2017). In addition, from 2017 a teaching guide was introduced to provide tips on optimisation of the training sessions and to ensure a more precise real-time assessment.

The resulting EUREP HOT format was verified yearly by collecting and prospectively analysing the following data: total number of participants attending the different HOT courses, participants' age, country of origin, and annual feedback obtained. Feedback was collected using Likert-scale questionnaires on tutor rating, course organisation, and personal expe- rience, with scores ranging from 1 (poor) to 5 (excellent). The scores were compared over time.

Results

A total of 796 participants from 54 countries participated in 1450 HOT sessions (response rate 89%) over the last 4 yr (Table 1). This included 294 (20%) ureteroscopy (URS) sessions, 237 (16.5%) transurethral resection (TUR) sessions, 840 (58%) basic laparoscopic (E-BLUS) sessions and 79 (5.5%) intermediate laparoscopic sessions. Participants' age ranged from 21 to 51 yr; the age demographics are shown in Table 1. Two-thirds of all participants (65.5%) were aged 30–34 yr, while 22% were aged 25–29 yr.

Regarding their country of origin, 712 participants (89%) were from Europe and 84 (11%) were from non-European nations. Of the European residents, most came from Italy (16%), Germany (15%), Spain (15%), and Romania (8%) (Table 2). Of the non-European nations, the majority came from India (4%) and Taiwan (2%).

The feedback for the E-BLUS session showed a constant improvement in score over the last 4 yr, with the highest scores achieved in 2017 (Table 3). This included feedback on improvements in tutor rating ($p = 0.017$), organisation ($p < 0.001$), and personal experiencewith EUREP ($p < 0.001$). While the overall rating for most aspects was between 4 (good) and 5 (excellent), overall the course duration had a slightly lower rating as most participants would like to have longer training sessions. However, this improved over the last 2 yr when the sessions were streamlined to 60-min sessions with a 1:1 trainee/tutor ratio. When comparing the scores from 2014 and 2017, there was a significant improvement in tutor rating ($p = 0.011$), organisation ($p = 0.005$), and personal experience with EUREP ($p < 0.001$).

Age range	Sessions, n (%)	TUR (n)	URS (n)	Laparoscopy (n)	
				Basic	Intermediate
20–24 yr	6 (0.5)	0	1	5	0
25–29 yr	297 (22)	37	76	172	12
30–34 yr	887 (65.5)	149	179	516	43
35–39 yr	147 (11)	26	25	81	15
40–45 yr	16 (1)	2	2	12	0
46–51 yr	1	1	0	0	0
Unknown	96	22	11	54	9

TUR = transurethral resection; URS = ureteroscopy.

Table 5: Age range of participants for the total of 1450 sessions performed during 2014-2017.

Participants (n)							
Europe				Asia		Africa	
Italy	112	Sweden	10	India	23	South Africa	3
Germany	110	Czech Republic	8	Taiwan	12	Egypt	3
Spain	106	Georgia	6	Lebanon	4	Algeria	2
Romania	56	Slovakia	6	Pakistan	2	**Unknown**	13
Poland	37	Belarus	5	Syria	1		
Turkey	32	Ukraine	5	Israel	1		
Portugal	26	Croatia	4	Hong Kong	1		
Belgium	23	Slovenia	4	China	1		
Greece	20	Serbia	**4**	**North America**			
Russia	17	Denmark	3	Canada	3		
France	18	Latvia	3	USA	1		
Switzerland	16	Macedonia	**3**	**South America**			
UK	14	Cyprus	2	Mexico	3		
Netherlands	13	Ireland	2	Brazil	2		
Hungary	12	Armenia	1	Colombia	1		
Lithuania	11	Azerbaijan	1	Paraguay	1		
Austria	10	Bulgaria	1	Peru	1		
Finland	10	Luxembourg	**1**	**Australia**	5	**Total**	591

Table 6: Participants from European and non-European countries during 2014-2017.

	2014	2015	2016	2017	*p* value	
					Trend analysis	2014–2017
Modifications to the EUREP training format						
Working hours per tutor (h)	22.1	22.7	24	22		
Trainee/tutor ratio	2:1	1:1	1:1	1:1		
Session duration (min)	Random	45	60	60		
Teaching guide available for tutors	No	No	No	Yes		
Residents undergoing basic laparoscopy HOT	183	236	225	196		
Evaluation by participants[a]						
How do you rate your tutor's teaching						
Focus on your individual training needs	4.7	4.7	4.6	4.8		
Providing helpful training advice	4.7	4.7	4.7	4.9		
Assessing your improvement	4.6	4.7	4.7	4.8		
Providing a structured training session	4.6	4.7	4.7	4.8		
Total	4.7	4.7	4.7	4.8	0.017	0.011
How do you rate the organisation						
Set-up of the equipment	4.7	4.7	4.7	4.7		
Duration of the course	3.9	3.7	4	4.1		
Number of tutors per trainee	4.7	4.8	4.8	4.9		
Time dedicated to the exercise	4.2	4	4.3	4.5		
Total	4.4	4.3	4.5	4.6	< 0.001	0.005
Personal experience						
Did it meet your expectations	4.3	4.4	4.4	4.6		
Did your basic laparoscopy skills improve	4.2	4.3	4.3	4.5		
Overall impression of the course	4.5	4.6	4.6	4.7		
Total	4.3	4.4	4.4	4.6	< 0.001	< 0.001

EUREP = European Urology Residents Education Programme; HOT = hands-on training.
[a] Scored as: 1 = bad; 2 = poor; 3 = fair; 4 = good; 5 = excellent.

Table 7: Feedback for EUREP basic laparoscopy HOT 2014-2017

Discussion

ESU started EUREP to harmonise and standardise training across Europe and to facilitate national societies in offer- ing curriculum-based training to residents. EUREP pro- vided a resident platform for simulation-based training and assessment in TUR, laparoscopy, and URS. HOT supplemented the theoretical knowledge gathered during this 5-d residential programme. The results and feedback demonstrate how valued it is for the residents involved. The individualised tutor ratio, organisation, and personal experience all add to evolution of the ideal course that other courses may strive to achieve. It is also a testament to the hard work of the ESU and the tutors involved, and their dedication, time, and effort in making the pro- gramme successful. As the focus of health care education changes, simulation has to be integrated into a comprehensive curriculum. Although these 1-h sessions will not make anyone an expert, the principle is to achieve incremental gains whereby all trainees have a chance to advance their competence and skill with 1:1 mentorship that builds on their previous knowledge and proficiency. This improvement is partly reflected in the feedback provided by the trainees of all age groups who attended EUREP from many European and non-European countries. Continuous evolution of EUREP is showcased by the adoption of new technology and bench models [8]. While providing top-quality training, EUREP tutors are also involved every year in the development of novel training protocols that, like E-BLUS and the more recent endoscopic stone treatment step 1 (EST-S1), are then translated to official European Association of Urology (EAU) assessment protocols ready for delivery even outside Europe. The tutors themselves need awareness of this to make the course more interesting and to keep it up to date [8,9]. An example is the use of K-box bench trainer for training in flexible URS [10].

While the course was structured and ran successfully for a number of years, owing to end-of-training HOT examinations such as E-BLUS and EST-S1, trainees who take these might feel nervous during the course itself [5,6]. However, the course also allows them to focus and prepare for the examination and maximise their potential for achievement. The other limitation of the EUREP format is difficulties in using more advanced training instruments such as lasers and in curricula with animal wet laboratory or cadaveric courses, although these could be done in other training centres in conjunction with EUREP. Standardised training protocols will allow more accurate and measurable training according to trainee needs. While basic models will be more useful to novices, more advanced models will allow com- plex and more realistic training [11,12] when needed. A combination of didactic teaching and supervised HOT not only improves surgical ability but is also helpful in real-life operating room environments [3].

The EUREP platform has allowed a successful era of E- BLUS and EST-S1 examinations [5,13] that can be a part of a trainee's portfolio and is now successfully conducted in numerous centres across the world. It seems that the EUREP journey has just begun, with new and more exciting courses and curricula on the horizon, such as the lower tract cur- riculum and advanced laparoscopic and stone treatment courses.

The feasibility of the EUREP HOT format, with 1:1 train- ing, 60-min time slots, and standardised teaching method- ology, is demonstrated by the enthusiasm and feedback given by the residents involved and reflects the annual improvements applied by the organisers. Moreover, this format allows easy planning of the training sessions and provides information on how many tutors are needed and the time needed to provide quality training to any target number of course participants. The gradual evolution of the course sets a format that other courses may strive to achieve, using EUREP as a benchmark for HOT courses. It is also a testament to the hard work of the ESU and the tutors involved for their dedication, time, and effort in making the programme successful.

Conclusions

The EUREP trainee demographics show that the purpose of the course is being achieved, with excellent feedback from the majority of trainees in the age group when they would be in their final year of training. While European trainees dominate the demographics, participation from a number of non-European countries suggests continued ESU collaboration with other national societies and wider dissemination of simulation training worldwide. The yearly improvements applied to the format have been well appreciated by participants. The EUREP HOT format, with its 1:1 standardised sessions of 60 min, is feasible andr eliable, which explains why the methodology is a solid base for EAU HOT courses around the world.

Author contributions: Bhaskar K. Somani had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Somani, Veneziano.
Acquisition of data: Van Cleynenbreugel, Gozen, Barmoshe, Biyani, Gaya, Hellawell, Pini, Rodriguez, Sanchez Salas, Macek, Skolarikos, Wagner, Eret, Haensel, Siena, Schmidt, Klitsch, Vesely, Ploumidis, Proietti, Kam- phuis, Tokas.
Analysis and interpretation of data: Veneziano.
Drafting of the manuscript: Somani.
Critical revision of the manuscript for important intellectual content:
Veneziano.
Statistical analysis: Geraghty.
Obtaining funding: None.
Administrative, technical, or material support: None.

Supervision: Palou, Van Cleynenbreugel, Gozen.

Other: None.

Financial disclosures: Bhaskar K. Somani certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultan- cies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

Funding/Support and role of the sponsor: EUREP is supported by educa- tional grant from Olympus.

Acknowledgments: We acknowledge all the EUREP hands-on-trainers and the EAU support staff.

References

1) Schwaitzberg SD, Connolly RJ, Sant GR, Reindollar R, Cleveland RJ. Planning, development, and execution of an international training program in laparoscopic surgery. Surg Laparosc Endosc 1996;6:10–5.
2) Geraghty R, Jones P, Somani BK. Worldwide trends of urinary stone disease treatment over the last two decades: a systematic review. J Endourol 2017;31:547–56.
3) Matsumoto ED, Hamstra SJ, Radomski SB, Cusimano MD. A novel approach to endourological training: training at the Surgical Skills Center. J Urol 2001;166:1261–6.
4) Fried GM. FLS assessment of competency using simulated laparoscopic tasks. J Gastrointest Surg 2008;12:210–2.
5) Brinkman WM, Tjiam IM, Schout BMA, et al. Results of the European basic laparoscopic urological skills examination. Eur Urol 2014;65:490–6.

6) Veneziano D, Ahmed K, Van Cleynenbreugel BV, et al. Development of the novel Endoscopic stone treatment step 1 (EST S1) training/ assessment curriculum: a collaborative work of EULIS, ESUT, YAUWP and ESU training groups. J Endourol 2017;31:934–41. http://dx.doi.org/10.1089/end.2017.0248.
7) Pietropaolo A, Proietti S, Geraghty R, et al. Trends of 'urolithiasis: interventions, simulation and laser technology' over the last 16 years (2000–2015) as published in the literature (PubMed): a systematic review. World J Urol 2017;35:1651–8.
8) Whitehurst LA, Somani BK. Semi-rigid ureteroscopy: indications, tips and tricks. Urolithiasis 2018;46:39–45. http://dx.doi.org/10. 1007/s00240-017-1025-7.
9) Chapman A, Somani BK, Robertson A, Kata SG. Decreasing cost of flexible ureterorenoscopy. Single use laser fibre cost analysis. Urol- ogy 2014;83:1003–5.
10) Villa L, Sener TE, Somani BK, et al. Initial content validation results of a new simulation model for flexible ureteroscopy: the Key-Box (K-BOX[1]). J Endourol 2017;31:72–7.
11) Sarmah P, Voss J, Veneziano D, Somani B. Low vs high fidelity: the importance of 'realism' in the simulation of a stone treatment procedure. Curr Opin Urol 2017;27:316–22.
12) Ahmed K, Jawad M, Abboudi M, et al. Effectiveness of procedural simulation in Urology: a systematic review. J Urol 2011;186:26–34.
13) Veneziano D, Minervini A, Beatty J, et al. Construct, content and face validity of the camera handling trainer (CHT): a new E-BLUS train- ing task for 30 laparoscope navigation skills. World J Urol 2016;34:479–84. http://dx.doi.org/10.1007/s00345-015-1657-6.

CHAPTER 2.3

Veneziano D, Ploumidis A, Proietti S, Tokas T, Kamphuis G, Tripepi G, Van Cleynenbreugel B, Gozen A, Breda A, Palou J, Sarica K, Liatsikos E, Ahmed K, Somani BK; ESU Training Group.

Validation of the endoscopic stone treatment step 1 (EST-s1): a novel EAU training and assessment tool for basic endoscopic stone treatment skills-a collaborative work by ESU, ESUT and EULIS.

World J Urol. 2019 Mar 27. doi: 10.1007/s00345-019-02736-4. Article of the Month May19

Validation of the Endoscopic Stone Treatment step 1 (EST-s1): A novel EAU training and assessment tool for basic endoscopic stone treatment skills - A collaborative work by ESU, ESUT and EULIS.

Domenico Veneziano[1,2,3,4], Achilles Ploumidis[5], Silvia Proietti[6], Theodoros Tokas[7], Guido Kamphuis[8], Giovanni Tripepi[9], Ben Van Cleynenbreugel[10], Ali Gozen[11], Alberto Breda[12], Joan Palou[12], Kemal Sarica[13], Evangelos Liatsikos[14], Kamran Ahmed[15], Bhaskar K Somani[16] and ESU training group.

1) Dept of Urology and Kidney Transplant, GOM, Reggio Calabria, Italy.
2) Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal.
3) ICVS/3B's, PT Government Associate Laboratory, Braga/Guimarães, Portugal
4) Dept. of Urology, Hofstra University School of Medicine, New York, NY, USA
5) Athens Medical Center, Athens, Greece
6) Dept. of Urology, San Raffaele Hospital, Ville Turro Division, Milan, Italy.
7) Department of Urology and Andrology, General Hospital Hall in Tirol, Austria
8) Academisch Medisch Centrum, Amsterdam, The Netherlands
9) Centro Nazionale Ricerca IFC, U.O. of Nephrology, Reggio Calabria, Italy
10) Department of Urology, University Hospital Gasthuisberg, Katholieke Universiteit Leuven, Louvain, Belgium
11) Department of Urology, Klinikum Heilbronn, Heilbronn, Germany
12) Department of Urology, Fundació Puigvert, Universitat Autònoma de Barcelona, Barcelona, Spain
13) Department of Urology, Dr. Lütfi Kģrdar Kartal Research and Training Hospital, Istanbul, Turkey
14) Department of Urology, University of Patras, Patras, Greece
15) Dept of Urology, Guy’s Hospital, London, United Kingdom
16) Department of Urology, University Hospital Southampton NHS Foundation Trust, Southampton, UK

Keywords: Stone-treatment; Training; EST-s1; Hands-on training; Education; curriculum; training-protocol.

Word count: 2964

Abstract

Introduction:

The Endoscopic Stone Treatment step 1 (EST s1) protocol has been developed after 2 years of collaborative work between different European Association of Urology (EAU) sections.

Objectives:

In this study we added construct validity evidence to the EST s1 curriculum.

Materials and Methods:

The EST-s1 curriculum includes 4 standardized tasks: flexible cystoscopy, rigid cystoscopy, semi-rigid URS and flexible URS.

Validation was performed during the annual 2016 EUREP meeting in Prague. 124 participants provided information on their endoscopic logbook and carried out these 4 tasks during a DVD recorded session. Recordings were anonymised and blindly-assessed independently by 5 proctors.

Inter-rater reliability was checked on a sample of 5 videos by the calculation of intra-class correlation coefficient. Task-specific clinical background of participants was correlated with their personal performance on the simulator. Breakpoint analysis was used to define the minimum number of performed-cases, in order to be considered “proficient”. “Proficient” and “Non-proficient” groups were compared for construct validity assessment. Likert-scale based questionnaires were used to test content and to comment on when the EST-s1 exams should be undertaken within the residency programme.

Results:

124 participants (105 final-year residents and 19 faculty members) took part in this study. The breakpoint analysis showed a significant change in performance curve at 36, 41, 67 and 206 seconds respectively corresponding to 30, 60, 25 and 120 clinical cases for each of the 4 tasks.

EST-s1 was scored as a valid training tool, correctly representing the procedures performed in each

task. Experts felt that this curriculum is best used during the third year of residency training.

Conclusion:

Our validation study successfully demonstrated correlation between clinical expertise and EST-s1 tasks, adding construct validity evidence to it. Our work also demonstrates the successful collaboration established within various EAU sections.

Patients summary:

In this study we validated the novel basic training curriculum for Endoscopic Stone Treatment, called EST-s1, developed after 2 years of collaborative work inside the EAU. We found that our curriculum is valid to certify proficiency of the basic endoscopic skills.

The protocol is now established as a standard tool to certify proficiency of the basic endoscopic skills in Europe, with potential of worldwide adoption in the future.

Take Home Message

Our study successfully added face and construct validity to EST-s1 tasks.

The protocol is now established as a standard tool to certify proficiency of the basic endoscopic skills in Europe, with potential of worldwide adoption in the future.

Introduction

The face of Endourology has changed over the last 2 decades, and with increasing use of ureteroscopy (URS) [1], a more structured and standardized training is required rather than the 'see one, do one, teach one' model. Despite the wide range of URS simulators available, standardised training is not yet possible due to the lack of a universally replicable curriculum [2,3].

We have previously described the development methodology of the European Association of Urology (EAU) novel Endoscopic Stone Treatment step 1 (EST-s1) training and assessment curriculum [4]. This was developed to have a standardised, low-cost, easily replicable set of exercises, to provide basic training and assessment within a framework of 45-60 minutes hands-on-training (HoT) sessions using bench training models. The curriculum was developed in collaboration with European School of Urology (ESU), Young Academic Urologists (YAU), European Section of Urolithiasis (EULIS), European section of Uro-Technology (ESUT) and European Urology Residents Education Programme (EUREP).

The development paper already added content validity evidence to the curriculum. The aim of this manuscript was to add face and construct validity evidence to the EAU EST-s1, to allow its regular use for basic endoscopic skills assessment of urologists in training.

Materials and Methods:

The EST-s1 curriculum development was carried out over 2 years in accordance with the EAU urolithiasis guidelines [5]. The development methodology started with a Cognitive Task Analysis [6,7] and followed the "full life-cycle curriculum development" [8] as described by Veneziano et al. [4]. The EST-s1 curriculum includes 4 tasks: 1)flexible cystoscopy; 2)rigid cystoscopy and safety guidewire placement; 3)semi-rigid ureteroscopy, placement of working guidewire and access sheath; 4)flexible ureteroscopy. Task goals and rules are summarized on a "tutor instructions sheet" (figure 1). Video explanation of all tasks can be found on the EAU website [9]. Validity evidence was considered in accordance with the Messick's framework [10], following the latest concepts in the field [11,12,13]. Test content, response process, internal structure and relationship to other variables were considered.

Data collection and methodology

Data collection was performed during the annual 2016 EUREP course in Prague. Equipment provided on each of the 5 stations available during the study is summarized by table 1. Participants enrolled in the study included final-year urology residents and members of the faculty. In order to objectively stratify expertise, disregarding the participation to the course as a resident or a faculty-member, each participant was asked to provide the number of procedures already performed, which directly related to the skills of the protocol: flexible and rigid cystoscopies, semi-rigid and flexible ureteroscopies. All participants carried out the 4 tasks during a single-participant session. They were well informed about the rules, which were explained by a proctor before the beginning of each session. Every participant had just one attempt for each task after a 1-minute warm up, intended to get used to "simulator anatomy". Each session was recorded on a DVD, which was then anonymised for blinded assessment. After the session, all participants were asked to fill-up a quality questionnaire to assess every exercise in relation to its “value as a basic training tool”, “the representation of the real counterpart”, “whether it should be an actual exercise of the final assessment protocol”. Scores provided in all questionnaires were Likert-scale based, ranging from 1 (“I don’t agree”) to 5 (“I fully agree”). Faculty members involved as participants were also asked to fill-out an additional questionnaire to contribute to *test content* [13] of each task. All participants were finally asked to comment about when the EST-s1 exams should be undertaken within the Urology residency programme.

Endoscopic Stone Treatment – Step 1 (EST – S1)

European School of Urology

Task 1: Flexible Cystoscopy (Target time 0:36 min:sec)

 Flexible cystoscope, guidewire (hard tip)	Enter the bladder and touch three of the marks, as indicated by the tutor. Guidewire has to be pre-loaded in the cystoscope. Guidewire has to be moved in/out to touch each target. **Marks:** fundus anterior wall left wall right wall tutor instructions are provided prior to time-count start.	**Time start/stop:** Scope enters the bladder (overcomes valve) —until— Guidewire touches the third mark **Errors:** incorrect use/position of the scope handle, incorrect response to tutors’ navigation requests advancing guidewire with bent cystoscope tip moving the guidewire sideways to touch the targets (without in/out movement)

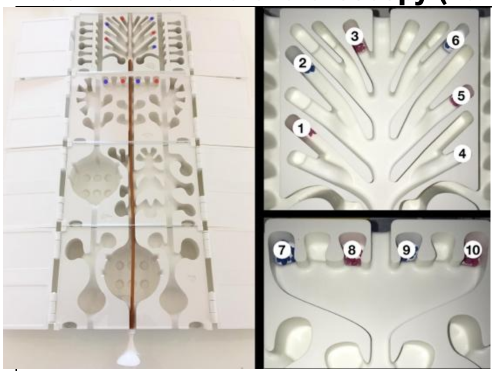
Task 2: Rigid Cystoscopy (Target times 0:41 min:sec)

 Rigid Cystoscope, safety guidewire	Assemble the cystoscope (optics, short bridge, sheath) Enter the bladder and touch two of the four marks, as indicated by the tutor. Guidewire has to be pre-loaded in the cystoscope. Guidewire has to be moved in/out to touch each target. Cannulate the ureter with the safety guidewire **Marks:** fundus anterior wall left wall right wall tutor instructions are provided prior to time-count start.	**Time start/stop:** Trainee touches the cystoscope —until— The trainee correctly assembles the cystoscope (max 1 minute) Scope enters the bladder (overcomes valve) —until— Scope exits the bladder (passing through the valve) leaving the guidewire in place **Errors:** incorrect assembly of the cystoscope excessive force application during rigid scope insertion not keeping urethral lumen central in the image during urethroscopy incorrect response to tutors’ navigation requests failing at cannulating the ureter with the guidewire

Task 3: Semi-rigid Ureteroscopy (Target time 1:07 min:sec)

 Semi-rigid Ureteroscope, working guidewire, access sheath	Perform a ureteroscopy next to the safety guidewire (already in place) and with the working guidewire on tip. Guidewire has to be pre-loaded in the ureteroscope. Leave the working guidewire in place Place the access sheath over the working guidewire.	**Time start/stop**: Scope enters the bladder (overcomes valve) —until— Exits the bladder with guidewire (safety and working) in place. **Errors:** Not keeping the ureteral lumen in the center of the field of vision for the majority of time. Failing at leaving working guidewire in place Placing the access sheath/guidewire with no lubrification Failing at positioning access sheath safely

Task 4: Flexible Ureteroscopy (Target time 3:26 min:sec)

 Flexible Ureteroscope, access sheath	Enter the access sheath with the flexible ureteroscope up to the most distal calices. Visualize and enter calices from 1 to 6 (touch the wall) and move back the access sheath to next proximal section ualize and enter calices from 7 to 10 (touch the wall) and exit the box with scope and sheath under direct vision.	**Time start/stop:** Scope enters access sheath —until— Scope and sheath come out of the box **Errors:** Rough handling of the scope Calices not correctly visualized Scope and sheath extracted not under direct vision.

General performance assessment

Depth perception (scale 1-5, pass: minimum 3)				
1. Constantly overshooting target, hits backstops, wide swings, slow to correct	.	**3.** Some overshooting or missing plane, but corrects quickly	.	**5.** Accurately directs instruments in correct plane to target
Bimanual dexterity (scale 1-5, pass: minimum 3)				
1. Use of one hand, ignoring non-dominant hand, poor coordination between hands	.	**3.** Use of both hands, but does not optimize interaction between hands to facilitate conduct of exercise	.	**5.** Expertly uses both hands in a mplementary manner to provide optimal working efficacy
Efficiency (scale 1-5, pass: minimum 3)				
1. Uncertain, much wasted effort, many tentative motions, constantly changing focus of exercise or persisting a task without progress	.	**3.** Slow but planned and reasonably organised	.	**5.** Confident, efficient and safe conduct of operation; maintaining focus on component of procedure until better done by another approach

Figure 9: tutor instruction sheet

Instrument/Tool	Serial Number
Flexible Video Ureteoscope (URF-V) 9.9 Fr. Insertion tube outer diam.	Olympus N30444760 / WA26780A
Ureterorenoscope URS, 6.4/7.8 Fr. 4.2 Fr. channel, 7° optics, angled ocular	Olympus WA29040B (with A0396 attachment)
Flexible cystoscope CYF-VH, 6.6 Fr. channel	Olympus N3828460 / N3828560
Cystoscope 4mm 30° rigid scope	Olympus A2012A
Cysto-Urethoscope Sheath 21 Fr.	Olympus A20913A
Albarran Double Channel Deflector	Olympus A20972A
Nitinol guidewire (Quattro)	Olympus EG50BX
Nitinol guidewire (EZ-Glider)	Olympus EG36BX
Access Sheath 24cm and 38cm	Olympus EG61224BX / EG61238BX
EndoUro trainer (tasks 1, 2, 3)	SAMED LS50
K-Box (task 4)	Porgès-Coloplast N/A

Table 8: Equipment provided

After the end of the EUREP event, DVDs were divided between 5 expert raters, with minimum 4 years of expertise in tutoring stone-treatment skills. A scoring sheet (figure 2) formulated during development process was provided to all raters. This included time-measurement along with "start-stop" information, quality criteria checklists and general performance assessment from the GOALS rating scale [14].

Inter-rater reliability was tested over 5 videos ranging from the fastest to the median and the slowest overall performance time to represent the spread across the participants enrolled. Time-measurement results collected for each task were compared to the number of clinical cases already performed as self-reported by each participant. The resulting curves were studied for *breakpoint analysis* [15]. The breakpoints identified on the curves were used to define the proficiency cut-off, intended as the number of previously performed cases needed to be considered "proficient". The "Quality assessment" and "General performance" assessments were also compared to previous clinical expertise of each participant, to check for any correlation.

Statistical analysis

Data were summarized as mean and standard deviation (SD) or as an absolute number, as appropriate. Inter-rater reliability between proctors was assessed by the calculation of intra-class correlation coefficient. Breakpoint analysis was used to highlight significant performance variations for each task analysed, in comparison with the number of cases performed by each participant. The selection of the best breakpoint and function type was based on maximizing the statistical coefficient of explanation and performing tests of significance [15]. The time-scores of participants who performed below and above the breakpoints for each task were compared by independent t-test. The inter-relationships among variables were assessed by Pearson product-moment correlation coefficient. Values of quality correspondent to the values of target times were derived by regression analysis. A P-Value <0.05 was considered statistically significant. All statistical analyses were done by SPSS for Windows (ver. 22), IBM, Chicago, Illinois, USA.

	Score Form EST Step 1	
Name of Examinee		**Name of Examiner**

Task 1 – Flexible cystoscopy **Time start:** scope enters the bladder **Time stop:** guidewire touches the third mark Guidewire pre-loaded in the cystoscope	**Trial 1**	**Trial 2** (only if trial 1 failed)
Time to complete task: To pass: **0.36** or less	(Min:sec)	(Min:sec)
Quality Criteria Scope correctly used and positioned	OK / Not OK	OK / Not OK
Marks (3) touched with guidewire as requested (within 1mm) (Critical)	OK / Not OK	OK / Not OK
Tutors' navigations requests carried out correctly	OK / Not OK	OK / Not OK
Guidewire advanced only with straight scope-tip	OK / Not OK	OK / Not OK
Task 2 – Rigid cystoscopy **Time start:** scope enters the bladder **Time stop:** guidewire enters the ureteral orifice Guidewire pre-loaded in the cystoscope	**Trial 1**	**Trial 2** (only if trial 1 failed)
Time to complete task: To pass: **0.41** or less	(Min:sec)	(Min:sec)
Quality Criteria Cystoscope correctly assembled in 1 minute **start: trainee touches the cystoscope** **stop: trainee correctly assembles the cystoscope**	OK / Not OK	OK / Not OK
Urethral lumen in the centre of the screen majority of time (during urethroscopy)	OK / Not OK	OK / Not OK
Tutors' navigations requests carried out correctly	OK / Not OK	OK / Not OK

Marks (2) touched with guidewire as requested (within 1mm) (Critical)	OK / Not OK	OK / Not OK
Ureteral orifice correctly cannulated with the guidewire (4 cm) (Critical)	OK / Not OK	OK / Not OK
Task 3 – Semi-rigid ureteroscopy **Time start:** scope enters the bladder **Time stop:** scope exits the bladder with guidewires (working and safety) in place Safety guidewire in place Working guidewire pre-loaded in the cystoscope **Access sheath placement not in time-count**	**Trial 1**	**Trial 2** (only if trial 1 failed)
Time to complete task: To pass: **1.07** or less	(Min:sec)	(Min:sec)
Quality Criteria Ureteral lumen in the centre of the screen majority of time (during ureteroscopy)	OK / Not OK	OK / Not OK
Working (second) guidewire successfully placed	OK / Not OK	OK / Not OK
Access sheath is wet and correctly assembled	OK / Not OK	OK / Not OK
Access sheath successfully inserted (Critical)	OK / Not OK	OK / Not OK
Task 4 – Flexible ureteroscopy **Time start:** scope enters the access sheath **Time stop:** scope comes out of the box along with the access sheath under direct vision	**Trial 1**	**Trial 2** (only if trial 1 failed)
Time to complete task: To pass: **3.26** or less	(Min:sec)	(Min:sec)
Quality Criteria Scope correctly used and positioned	OK / Not OK	OK / Not OK
Calices from 1 to 6 visualized correctly with the tip touching the number (Critical)	OK / Not OK	OK / Not OK
Calices from 7 to 10 visualized correctly with the tip touching the number (Critical)	OK / Not OK	OK / Not OK

Scope and access sheath removed safely and under direct vision (Critical)	OK / Not OK	OK / Not OK

Global performance assessment

Depth perception (scale 1-5, pass: minimum 4)				
1. Constantly overshooting target, hits backstops, wide swings, slow to correct	**2.**	**3.** Some overshooting or missing plane, but corrects quickly	**4.**	**5.** Accurately directs instruments in correct plane to target
Bimanual dexterity (scale 1-5, pass: minimum 3)				
1. Use of one hand, ignoring non-dominant hand, poor coordination between hands	**2.**	**3.** Use of both hands, but does not optimize interaction between hands to facilitate conduct of exercise	**4.**	**5.** Expertly uses both hands in a complementary manner to provide optimal working exposure
Efficiency (scale 1-5, pass: minimum 3)				
1. Uncertain, much wasted effort, many tentative motions, constantly changing focus of exercise or persisting a task without progress	**2.**	**3.** Slow but planned and reasonably organised	**4.**	**5.** Confident, efficient and safe conduct of operation maintaining focus on component of procedure until better done by another approach

Figure 10: scoring sheet

Results

Overall, 124 participants were enrolled in the study. 105 of them were final-year residents, while 19 were part of the faculty. Within the faculty members who took part in the study, 10 performed >10 URS procedures/month (avg. 23.7/month, SD± 13,5) and were considered for *content validity* re-assessment. Resident demographics are reported in table 2. Inter-relationship between variables showed agreement by 99.9% regarding time-measurement and general performance scores among the 5 proctors involved. The 4 tasks were analysed separately in order to obtain more detailed validation results. Therefore, results from the different tasks will be presented in separate paragraphs. All data are summarized in table 9.

Age 26-28 years – 12 (11.4%) 29-30 years – 52 (49.5%) 31-33 years – 36 (34.2%) No answer – 5 (4.7%)
Countries European/Non-Europeans (%) 95.3% vs 4.7% Spain (n=29), Italy (n=22), Romania (n=9), Germany (n=9), Belgium (n=6), Poland (n=6), Portugal (n=5), Hungary (n=4), Finland (n=4), Greece (n=3), The Netherlands (n=2), Others (n=5)

Table 9: Trainee demographics and feedback from participants

	Task 1 flex cystoscopy	Task 2 rigid cystoscopy	Task 3 semi-rigid URS	Task 4 flex URS	
Clinical background					**N° of cases**
Residents avg.	209 (±428)	190 (±324)	47 (±77)	16 (±30)	
Residents min - max	0 - 3000	0 - 3000	0 - 600	0 - 200	
Faculty members avg.	1384 (±974)	1323 (±924)	781 (±719)	781 (±556)	
Faculty members min - max	100 - 3000	100 - 3000	20 - 2000	20 - 2000	
Time assessment					**Construct**
performed below the breakpoint	50%	34%	49%	19%	
proficient outperformed the rest	67%	58%	45%	48%	
Quality criteria					
correlation with proficiency	1 on 4	0 on 4	0 on 5	1 on 4	
General performance assessment					
correlation with proficiency	3 on 3	3 on 3	3 on 3	3 on 3	
Feedback (avg out of 5)					**Face**
It's a valid training tool	4.6	4.6	4.6	4.7	
Correctly represents the procedure	4.5	4.7	4.3	4.4	
It should be part of EST s1	4.5	4.7	4.7	4.6	

Table 10: data summary

The 105 residents enrolled reported the following average number of procedures already performed: 209 flexible cystoscopies (SD±427.9, Min 0, Max 3000), 190 rigid cystoscopies (SD±324, Min 0, Max 3000), 47 semi-rigid ureteroscopies (SD±76,9, Min 0, Max 600) and 16 flexible ureteroscopies (SD±30, Min 0, Max 200). The 19 faculty members enrolled reported the following average number of procedures already performed: 1384 flexible cystoscopies (SD±974, Min 100, Max 3000), 1323 rigid cystoscopies (SD±924, Min 40, Max 3000), 781 semi-rigid ureteroscopies (SD±719, Min 20, Max 2000), 461 flexible ureteroscopies (SD±556, Min 20, Max 2000).

Breakpoint analysis for the collected results on task 1 showed a significant change in the performance curve at 36 seconds, corresponding to 30 flexible cystoscopies performed. 50% of the participants (n=62) had already performed at least 30 flexible cystoscopies before the trial and were considered part of the “proficient” group. This group outperformed the rest of the participants by 67%. Breakpoint analysis for the collected results on task 2 showed a significant change in the performance curve at 41 seconds (P<0.05), corresponding to 60 rigid cystoscopies performed. 34% of the participants (n=42) had already performed at least 60 rigid cystoscopies before the trial and were considered part of the “proficient” group. This group outperformed the rest of the participants by 58%. Breakpoint analysis for the collected results on task 3 showed a significant change in the performance curve at 67 seconds (P<0.05), corresponding to 25 semi-rigid ureteroscopies performed. 49% of the participants (n=61) had already performed at least 25 semi-rigid ureteroscopies before the trial and were considered part of the “proficient” group. This group outperformed the rest of the participants by 45%. Breakpoint analysis for the collected results on task 4 showed a significant change in the performance curve at 206 seconds (P<0.05), corresponding to 120 cases performed. 19% of the participants (n=24) had already performed at least 120 flexible ureteroscopies before the trial and were considered part of the “proficient” group. This group outperformed the rest of the participants by 48% (table 11).

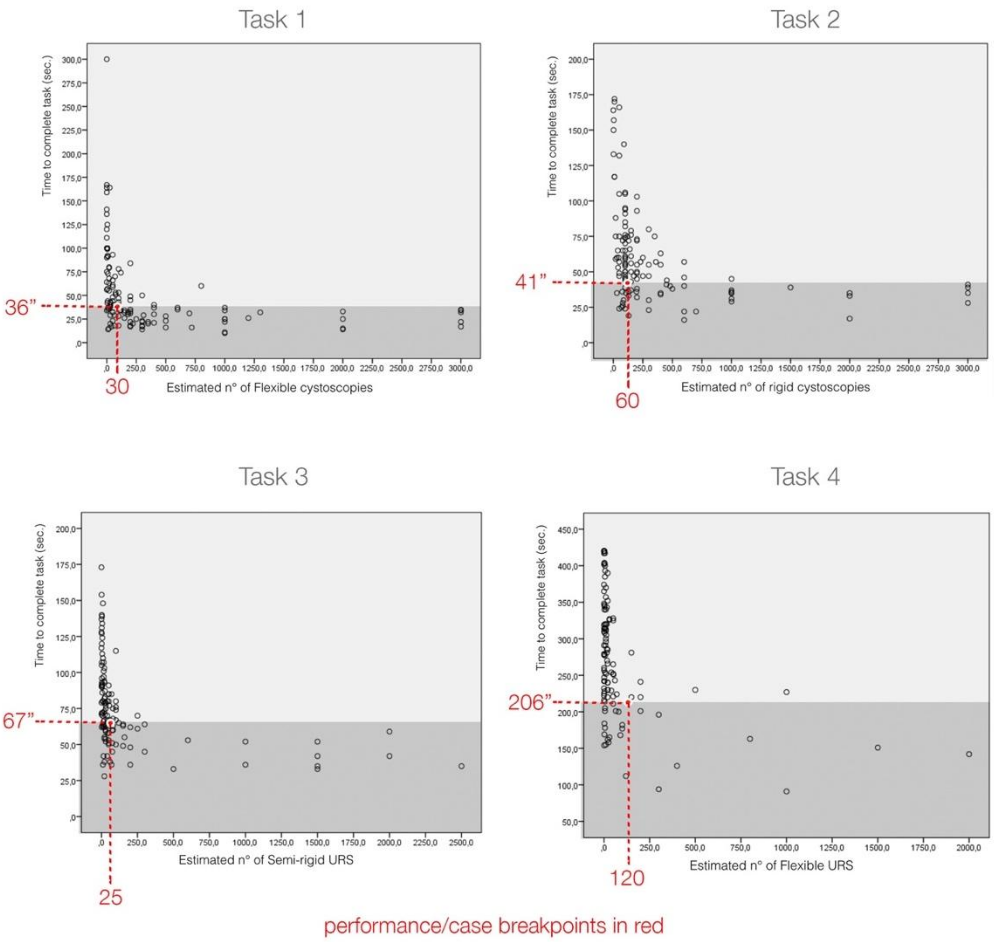


Table 11: breakpoint analysis

Pearson correlation between the quality-assessment checklists and the number of procedures already performed is summarized in table 3. The correlation was found to be statistically significant only for task 1/quality item 3, "Tutor's navigations requests carried out correctly" (P=0.04) and for task 4/ quality item number 4, "Scope and access sheath removed safely and under direct vision" (P=0.05).
Pearson correlation between the 3 General performance assessment items and the number of procedures already performed is summarized in table 2. The correlation was found to be statistically significant for all tasks in relation to all the items ($P \leq 0.0001$).
The 10 selected faculty members provided the following average scores on their questionnaires. "It is a valid training tool" scored 4.6 for Task 1, 4.5 for Task 2, 4.6 for Task 3, 4.7 for Task 4. "Correctly represents the corresponding procedure" scored 4.5 for Task 1, 4.7 for Task 2, 4.3 for Task 3, 4.4 for Task 4. "Should be part of the EST s1" scored 4.7 for Task 1, 4.7 for Task 2, 4.7 for Task 3, 4.6 for Task 4.

Common data

The average scores collected by the 105 residents enrolled were the following: "it is a valid training tool",4.6; "correctly represents a URS procedure", 4.2 and "should be used for basic URS training & assessment", 4.5. For the question "when would you suggest the ESTs1 training along the 5 years of residency" the mean was 3.4, while the mode was 5 (31,5% of voters). The 10 faculty members selected for the high volumes of clinical procedures performed, considered that the ESTs1 "should be used for basic URS training & assessment" with a mode of 5 (80% of all voters), while suggested it to be incorporated as a training and assessment curriculum during the 3rd year of residency (Mode 3, 40% of all voters).

Discussion

Meaning of our study:

EST-s1 is developed as the first of a three-step curriculum (table 5) and is designed to replicate the basic skills of an endoscopic stone-treatment procedure. This protocol is based on low-fidelity simulators [16] and contains skills that are common to several procedures. Indeed, the EST-s1 protocol can be easily adopted in Gynaecology, Gastroenterology or General Surgery to teach or assess the use of rigid and flexible scopes. The same approach has been applied in the last 5 years to laparoscopy with relevant results [17,18,19,20].

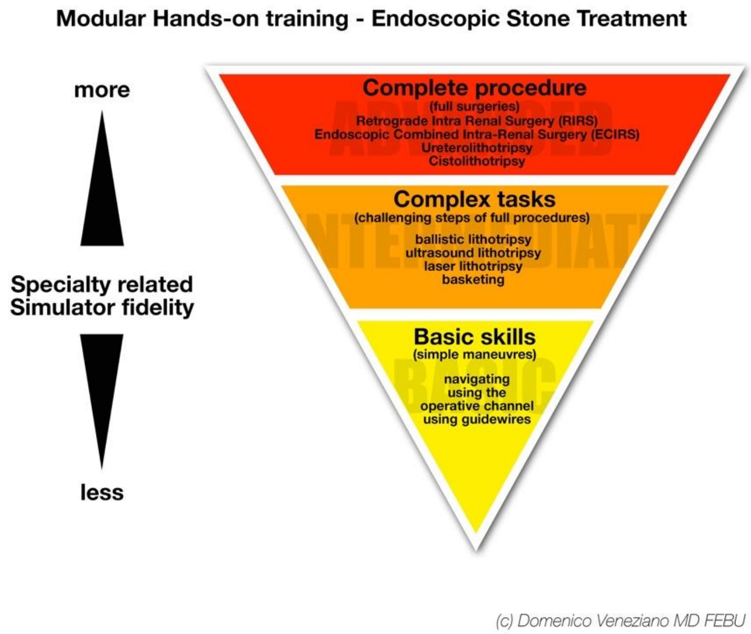


Table 12: Modular hands-on training template for EST full curriculum

Having previously described the validity of content [4], our study was aimed to adding construct validity evidence to the novel EST-s1 protocol, following the latest concepts in the surgical education field [13]. Relevant data will be analysed below, in consideration of our main aims.

The 5 raters involved in the blinded assessment actively participated in the development of the protocol. They had at least 4 years of expertise as tutors for stone-treatment skills and were well informed about the rules, which ensured that the *response process* requirements were fulfilled.

Internal structure of the analysis process was confirmed by the inter-rater reliability assessment and by the blinded review of all recorded sessions, operated by expert raters, in line with the latest evidences for curriculum validation. The raters scored additional 5 videos for “Time-measurement” and “General performance”. The results collected showed inter-rater reliability by 99.9% of agreement.

The correlation with the previous clinical expertise and the performance gap recorded on the simulator between proficient and non-proficient participants allowed us to demonstrate *relationship to other variables* for each task. Considering this correlation allowed us to avoid one of the most common biases: the “arbitrary” division between *Experts* and *Novices.* The cut-offs are usually defined without clear reasons and can be a bias for validation studies. Therefore, before comparing the “proficient” group with the “non-proficient” group, we used breakpoint analysis to define the correct proficiency cut-off for each task, in consideration of the individual number of cases already performed and its correlation with simulation performance. The data collected showed that the decision taken was paramount: for the basic skills tested, being a resident did not always correspond to being a "novice", while being a faculty member was not always equivalent to be an "expert" as the number of procedures performed by a resident was not always low and vice-versa. For example, in relation to task 3 (semi-rigid ureteroscopy), the minimum number of procedures previously performed within the faculty group was 20, while some residents had already performed 600 semi-rigid ureteroscopies. Following this concept, clinical background directly related to each single task was used to stratify the 124 participants involved and this data was put into correlation with their per-task performance. The breakpoint analysis provided us with the cut-off number of cases needed to be considered part of the "proficient" group. This group outperformed the "non-proficient" one at least by 45% on each single task.

After the session, the residents enrolled were provided with questionnaires for *test content* re-assessment of the entire protocol. According to the scores collected among the residents enrolled, the protocol scored above the average of 4/5 (agree) in relation to all the statements suggested, meaning that it was considered a "valid training tool", which "correctly represents a URS procedure" and "should be used for basic URS training & assessment". The faculty members confirmed the last statement with a mode of 5. According to the faculty and in consideration of the basic contents proposed, it was considered as an appropriate training tool for 3rd year residents, which was also the average residency-year voted by residents. Despite the differences between the several residency curriculums already established all over Europe, this evaluation might work as a suggestion for introducing basic endoscopic skills as part of the residency pathway, providing more information on how and when putting EST-S1 curriculum into practice.

Correlation between the quality assessment checklist and actual proficiency of the participants was poor. It was already demonstrated by Regehr et al. [21] that global rating scales like the GOALS are superior to checklists specifically for the evaluation of technical skills, especially when expert raters were involved. Despite our attempt to integrate the GOALS with further information, our study confirmed these findings. Quality assessment items will be in any case preserved in the final examination sheet and their functionality will be re-evaluated after the collection of a higher data volume.

The trainees enrolled for our study came from several European countries and this might be considered as a limitation in our paper as mixed-countries participants' cohort could correlate to a variable clinical background in terms of knowledge of procedural steps. This means that our protocol might not fit the 3rd year of residency in every geographical area, depending on the different residency programs. Moreover, being the sample mostly limited to the European experience (95.3% of participants), cut-offs might not exactly apply to different environments, which could need further investigation with samples from other continents. Another pitfall could be related to self-reported clinical background data. Hard metrics extracted from medical records or direct measurements, in place of self-reported counterparts might have solved this problem, but might be difficult to collect from a large cohort of participants. The same issue was already reported by other validation papers [20]. Finally, integrating a cognitive on-line part and non-technical skills [22,23] to our current training curriculum could also add further value to it.

Conclusion

We successfully demonstrated correlation between clinical expertise and EST-s1 tasks. Passing EST-s1 tasks certifies proficient mastery of the basic endoscopic skills. The protocol is now established as a standard training and assessment tool for Endoscopic skills in Europe with a potential for worldwide adoption in future. Our work also demonstrates the successful collaboration established within various EAU sections.

Author's contribution

D Veneziano: Project development, Data Collection, Manuscript writing

A Ploumidis: Data Collection

S Proietti: Data Collection

T Tokas: Data Collection

G Kamphuis: Data Collection

G Tripepi: Statistics

B Van Cleynenbreugel: Data Collection

A Gozen: Data Collection

A Breda: Data Collection

J Palou: Data Collection

K Sarica: Data Collection

E Liatsikos: Data Collection

K Ahmed: Data Collection

BK Somani: Project development, Manuscript writing

Disclosure of potential conflicts of interest:

The authors declare that they have no conflict of interest.

Research involving Human Participants and/or Animals

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study.

References

1. Geraghty R, Jones P, Somani BK. Worldwide Trends of Urinary Stone Disease Treatment over the last two Decades: A Systematic Review. J Endourol, 2017 Jun;31(6):547-556.
2. Brunckhorst O, Aydin A, Abboudi H, et al. Simulation-based ureteroscopy training: A systematic review. J Surg Educ. 2015 Jan-Feb;72(1):135-43.
3. Seitz C., Fajkovic H., Training in ureteroscopy for urolithiasis. Arab J. Urol. 2014 Mar; 12(1): 42-8. Doi: 10.1016/ j.aju.2013.08.010.
4. Veneziano D, Ahmed K, Van Cleynenbreugel B, Gozen A, Palou J, Sarica K, Liatsikos E. Development methodology of the novel Endoscopic stone treatment step 1 (EST s1) training/assessment curriculum. J Endourol, END-2017-0248-OR.R1
5. Türk C, Neisius A, Petrik A, Seitz C, Skolarikos A, Tepeler A, Thomas K Guidelines Associates: Dabestani S, Drake T, Grivas N, Ruhayel Y. 2014 publication of the EAU Urolithiasis Guidelines.
6. D. Hananel JS, R. Sweet. Simulator Development – from Idea to Prototype to Product. In: R. Aggarwal JKJC-B, editor. ACS Principles and Practice for Simulation and Surgical Education Research. Chicago, IL: American College of Surgeons; 2015.
7. Brinkman W, Tjiam IM, Schout BMA, Hendrikx AJM, Witjes JA, Scherpbier AJJA, et al. Designing simulator-based training for nephrostomy procedure: an integrated approach of

cognitive task analysis (CTA) and 4-component instructional design (4C/ID). Journal of Endourology. 2011;25:A29-A.

8. Khamis NN, Satava RM, Alnassar SA, Kern DE. A stepwise model for simulation-based curriculum development for clinical skills, a modification of the six-step approach. Surgical Endoscopy and Other Interventional Techniques. 2016;30(1):279-87.
9. EST s1 video instructions: http://uroweb.org/education/online-education/surgicaleducation/endoscopic-stone-treatment-est/
10. Messick S. Validity of psycological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. 1994. https://doi.org/10.1002/j.2333-8504.1994.tb01618.x
11. Korndorffer JR, Kasten SJ, Downing SM. A call for the utilization of consensus standards in the surgical education literature. Am J Surg. 2010;199(1):99-104.
12. Sweet RM, Hananel D, Lawrenz F. A unified approach to validation, reliability , and education design for for surgical technical skills training.. Arch Surg. 2010; 145(2):197-201.
13. Goldenberg M, Lee JY. Surgical Education, Simulation, and Simulators – Updating the Concept of Validity. Curr Urol Rep (2018) 19: 52
14. Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondrè K, Stanbridge D, Fried GM. A global assessment tool for evaluation of intraoperative laparoscopic skills. Am J Surg 190 (2005) 107-113
15. Jones RH, Molitoris BA. A statistical method for determining the breakpoint of two lines. Anal Biochem 1984 Aug 15;141 (1): 287-90.
16. Sarmah P, Voss J, Ho A, Veneziano D, Somani B. Low vs high fidelity: the importance of 'realism' in the simulation of a stone treatment procedure. Curr Opin Urol 2017.
17. Brinkman WM, Tjiam IM, Schout BMA, Muijtjens AMM, Van Cleynenbreugel B, Koldewijn EL, et al. Results of the European Basic Laparoscopic Urological Skills Examination. Eur Urol. 2014;65(2):490-6.
18. Fried GM. FLS assessment of competency using simulated laparoscopic tasks. Journal of Gastrointestinal Surgery. 2008;12(2):210-2. World J Urol - DOI 10.1007/s00345-015-1657-6
19. Veneziano D, Minervini A, Beatty J, et al. Construct, content and face validity of the camera handling trainer (CHT): a new E-BLUS training task for 30° laparoscope navigation skills. World J Urol. 2016 Apr;34(4):479-84. doi: 10.1007/s00345-015-1657-6. Epub 2015 Aug 6.

20. Kowalewski TM, Sweet R, Lendvay TS, et al. Validation of the AUA BLUS tasks. J Urol. 2015, Vol. 195, 998-1005
21. Regehr G, MacRae H, Reznick RK, et al. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. Acad Med 1998;73:993-7
22. Wood TC, Raison N, Haldar S, et al.: Training Tools for Nontechnical Skills for Surgeons—A Systematic Review. J. Surg. Educ. 2016, doi:10.1016/j.jsurg.2016.11.017.
23. Brunckhorst O, Shadid S, Aydin A, Khan S, McIlhenny C, Brewin J. The Relationship Between Technical And Nontechnical Skills Within A Simulation-Based Ureteroscopy Training Environment – J Surg Educ. Sept-Oct, 2015. Vol 72, Issue 5, Pages 1039-1044

CHAPTER 2.4

Veneziano D, Ploumidis A, Proietti S, Tokas T, Kamphuis G, Tripepi G, Van Cleynenbreugel B, Gozen A, Breda A, Palou J, Sarica K, Liatsikos E, Ahmed K,

Somani BK, on behalf of the European School of Urology training group

Evolution and Uptake of the Endoscopic Stone Treatment Step 1 (EST-s1) protocol: establishment, validation and assessment in a collaboration by the European School of Urology and the Uro-Technology and Urolithiasis Sections.

Eur Urol. 2018, doi:10.1016/j.eururo.2018.05.012

The evolution and uptake of Endoscopic Stone Treatment step 1 (EST-s1): Protocol establishment, validation and assessment from a collaborative work by ESU, ESUT and EULIS.

Domenico Veneziano [a,b,c,*], Achilles Ploumidis [d], Silvia Proietti [e], Theodoros Tokas [f], Guido Kamphuis [g], Giovanni Tripepi [h], Ben Van Cleynenbreugel [i], Ali Gozen [j], Alberto Breda [k], Joan Palou [k], Kemal Sarica [l], Evangelos Liatsikos [m], Kamran Ahmed [n], Bhaskar K. Somani [o] on behalf of the European School of Urology training group

A) Department of Urology and Kidney Transplantation, Ospedali Riuniti BMM, Reggio Calabria, Italy

B) Life and Health Sciences Research Institute, School of Medicine, University of Minho, Braga, Portugal

C) PT Government Associate Laboratory, Life and Health Sciences Research Institute, Braga, Portugal

D) Athens Medical Centre, Athens, Greece

E) Department of Urology, San Raffaele Hospital, Milan, Italy

F) Department of Urology and Andrology, General Hospital, Hall in Tirol, Austria

G) Academisch Medisch Centrum, Amsterdam, The Netherlands

H) Nephrology Operative Unit, Centro Nazionale Ricerca IFC, Reggio Calabria, Italy

I) Department of Urology, University Hospital Gasthuisberg, Katholieke Universiteit Leuven, Louvain, Belgium

J) Department of Urology, Klinikum Heilbronn, Heilbronn, Germany

K) Department of Urology, Fundació Puigvert, Universitat Autònoma de Barcelona, Barcelona, Spain

L) Department of Urology, Dr. Lütfi Kgrdar Kartal Research and Training Hospital, Istanbul, Turkey

M) Department of Urology, University of Patras, Patras, Greece

N) Department of Urology, Guy's Hospital, London, UK

O) Department of Urology, University Hospital Southampton NHS Foundation Trust, Southampton, UK

*Corresponding author. Department of Urology and Kidney Transplantation, Ospedali Riuniti BMM, Via Giudecca 52, Reggio Calabria 89125, Italy. Tel. +39 32 0479 7143. E-mail address: info@domenicoveneziano.it (D. Veneziano).

Word count: 498 words
Key words: Ureteroscopy, Simulation, training, endoscopy, endourology

Endourology training has evolved over the last two decades with more emphasis now being placed on simulation-based training. While EBLUS training-curriculum and examination have been well established (1), there was a lack of standardised training for endourology. The European School of Urology (ESU), together with the European Section of Uro-Technology (ESUT) and the European Section of Urolithiasis (EULIS), started the development of the Endoscopic Stone Treatment step-1 (EST-s1) simulation protocol in 2014. This was produced in accordance with the EAU Guidelines by following the "full life-cycle curriculum development" template. The outcomes and metrics were defined through a cognitive task analysis via the EAU Young Academic Urology (YAU) group and furthermore the simulator requirements were tested. The final task list consisted of four exercises which replicated the basic skills of endoscopic stone treatment: 1)flexible cystoscopy 2) rigid cystoscopy and placement of a safety guidewire 3)semi-rigid ureteroscopy and access-sheath placement 4) flexible uretero-renoscopy.

Curriculum development process took 2 years of consensus meetings and expert consultation which lead to adding content validity evidence to the protocol (2). Face and construct validity data collection happened during the annual EUREP course in 2016 (3). This validation study involved 124 participants using low-fidelity simulators. The rules of the exercises and expected goals were strictly derived from the development process and were summarised in a "tutor instruction sheet". Video explanation of tasks was available on the EAU website. A good correlation was seen between the EST-s1 assessment and clinical background.

Once the EST-s1 curriculum was validated, it was then adopted for trainee assessment and certification for the first time during EUREP17. The examination was conducted with trainees getting a 1-minute warm up at the beginning of the session, to get used to the "functionality" of the simulator. For certification, they had to successfully complete all the tasks with no errors and below the cut-off time given for each task, also taking into account their "global performance assessment", as rated by the tutor. Participants had maximum two attempts to pass each exercise, with a 1:1 trainee:tutor set-up and under tutors' instructions.

EST-s1 examinations are now established and have been carried out in the annual EAU, EULIS, ESUT and EUREP meetings over the last year. Feedback from the various changes adopted in the training curriculum has been excellent (4) and over 20 EST-s1 training requests are pending from several National Urological Societies around the world, which expressed the intention to officially adopt the protocol during their events.

EST-s1 is planned to be the first part of a 3-stepped curriculum designed to bring standardisation in the field of Endourology training. It is a low-cost, replicable and validated training protocol, with a strong potential for further European and worldwide adoption; the first to be fully developed within the EAU, in accordance with EAU evidence-based clinical standards. Its development demonstrated possible successful collaborations within EAU sections, aimed to the establishment of robust and shared training protocols worldwide with the final goal of avoiding the use of patients as training platforms.

References:

[1] Brinkman WM, Tjiam IM, Schout BMA, et al. Results of the European basic laparoscopic urological skills examination. Eur Urol 2014;65:490–6.

[2] Veneziano D, Ahmed K, Van Cleynenbreugel B V, et al. Development of the novel Endoscopic Stone Treatment step 1 (EST s1) training/ assessment curriculum: an international collaborative work by European Association of Urology sections. J Endourol. In press. https://doi.org/10.1089/end.2017.0248.

[3] Somani BK, Van Cleyenbreugel B, Gozen A, et al. The European Urology Residents Education Programme hands-on-training for- mat: 4 years of hands-on training improvements from the European School of Urology. Eur Urol Focus. In press. https://doi.org/10.1016/j. euf.2018.03.002.

CHAPTER 2.5

Veneziano D, Canova A, Arnolds M, Beatty JD, Biyani SC, Deho F, Fiori C, Hellawell G, Langenhuijsen H, Pini G, Rodriguez Faba O, Siena G, Skolarikos A, Tokas T, Van Cleynenbreugel B, Wagner C, Tripepi G, Somani B, Lima E.

The Performance Improvement (Pi) score: an algorithm to objectively score performance improvement during E-BLUS hands on training (HoT) sessions. An ESUT project.

BJU Int. 2018 Nov 15. doi: 10.1111/bju.14621.

Performance Improvement (Pi) score: an algorithm to score Pi objectively during E-BLUS hands-on training sessions. A European Association of Urology, Section of Uro-Technology (ESUT) project

Domenico Veneziano[1,2,3], Antonio Canova[4], Michiel Arnolds[5], John D. Beatty[6], Chandra S. Biyani[7], Federico Deho[8,9], Cristian Fiori[10], Giles O. Hellawell[11], J.F. Langenhuijsen[12], Giovannalberto Pini, Oscar Rodriguez Faba[14], Giampaolo Siena[15], Andreas Skolarikos[16], Theodoros Tokas[17], Ben S.E.P. Van Cleynenbreugel[18], Christian Wagner[19], Giovanni Tripepi[20], Bhaskar Somani[21] and Estevao Lima[1,2]

1) Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga,
2) ICVS/3B's, PT Government Associate Laboratory, Braga/ Guimarães, Portugal,
3) Department of Urology and Kidney Transplant, Grande Ospedale Metropolitano, Reggio Calabria,
4) Human Resources Management, MBDA Italia SpA, Rome, Italy,
5) Department of Urology, Gelre Ziekenhuis, Apeldoorn, The Netherlands,
6) Department of Urology, Leicester General Hospital, University of Leicester, Leicester,
7) Department of Urology, St James's University Hospital Leeds Teaching Hospitals NHS Trust, Leeds, UK,
8) Department of Urology, San Raffaele Scientific Institute, University Vita-Salute,
9) Division of Oncology/Unit of Urology, URI, IRCCS Ospedale San Raffaele, Milan,
10) Department of Oncology, A.O.U. San Luigi Gonzaga, Università degli Studi di Torino, Orbassano, Italy,
11) North West London Hospitals NHS Trust, Imperial College Healthcare NHS Trust, London, UK,
12) Department of Urology, Radboud University Medical Centre, Nijmegen, The Netherlands,
13) Department of Urology, Ospedale San Raffaele – Turro, Milano, Italy,
14) Department of Urology, Fundació Puigvert, Barcelona, Spain,
15) Department of Urology, Azienda Ospedaliero- Universitaria di Careggi, Firenze, Italy,
16) Department of Urology, University of Athens, Athens, Greece,
17) Department of Urology and Andrology, General Hospital Hall i.T., Hall in Tirol, Austria,
18) Department of Urology, UZ Leuven, Leuven, Belgium,
19) Department of Urology and Oncology, St. Antonius-Hospital, Gronau, Germany,
20) CNR IFC, U.O. of Reggio Calabria, Reggio Calabria, Italy, and

21) Department of Urology, University Hospital Southampton NHS Foundation , Southampton, UK

Abstract

Introduction and objectives

Assessing performance improvement (Pi) is one of the most important roles of a tutor, but is usually based on subjective observation, personal judgement and expertise. Our study is aimed to evaluate the variability of subjective tutor Pi assessment and to compare it to a novel measurement algorithm: 'the Pi-score'

Materials and methods

The Pi-score algorithm considers time measurement and number of errors from two different repetitions (first and fifth) of the same training task and compares them to the relative task goals, to produce an objective score. We collected data during 8 courses on the four basic laparoscopic urological skills (E-BLUS) tasks. The same tutor instructed on all courses. Collected data were independently analysed by 14 hands-on-training (HoT) experts for Pi assessment. Their *subjective* Pi assessments were compared for inter-rater reliability. The average per-participant *subjective* scores from all 14 proctors were then compared to *objective* Pi-score algorithm results. Cohen's Kappa Statistics was used for comparison analysis.

Results

50 participants were enrolled. Concordance found between the 14 proctors' scores was the following: Task1=0.42 (moderate); Task2=0.27 (fair); Task3=0.32 (fair); Task4 (Kappa=0.55, moderate). Concordance between Pi-score results and proctor average scores per participant was the following: Task1=0.85 ("almost perfect"); Task2=0.46 ("moderate"); Task3=0.92 ("almost perfect"); Task4=0.65 ("substantial").

Conclusion

Our study demonstrated that evaluation of Performance improvement is highly variable, even when formulated by a cohort of experts. Our algorithm successfully provided an objective score that equals the

average Pi assessment of a cohort of experts, in relation to a small amount of training attempts.

Introduction

Performing surgical training outside the operative room (OR) is considered paramount in order to prepare for 'real life' practice [1]. Finally, the aim is usually focussed on surgical skills' details, improving performance and, most important, avoiding patients to be used as training platforms. Surgical training, as stated by R. Satava [2], relies mostly on the curriculum, rather than on the simulator used, which means that the tutor plays a critical role in it.

Analysing the improvement of a student is usually one of the most important roles of the Hands-on Training (HoT) tutors. They observe the trainees along their training trials and measure the achievements, in order to judge the improvement of performance along the HoT session. However, this process is not objective or standardized and relies on personal expertise. Many attempts have been made to objectively measure performance improvement of a student, in relation to surgical practical skills. Indeed, measuring performance improvement can be very useful to develop more reliable training methodologies and optimize any educational procedure. Data acquired are usually put into a sequence, to trace a "learning-curve"[3] and these have also been also analysed to study innate ability [4]. Learning curve sometimes considers just one parameter, usually the time-to-complete a task, without matching it to other critical data, as for example the "number of errors" or the "time-goal" for that specific task. In other instances, learning curve connects a series of scores given by a simulator to the single task performed [5]. This allows performance improvement measurement and weighted error assessment, but just on virtual reality platforms or through expensive data collection systems like the EDGE [6].

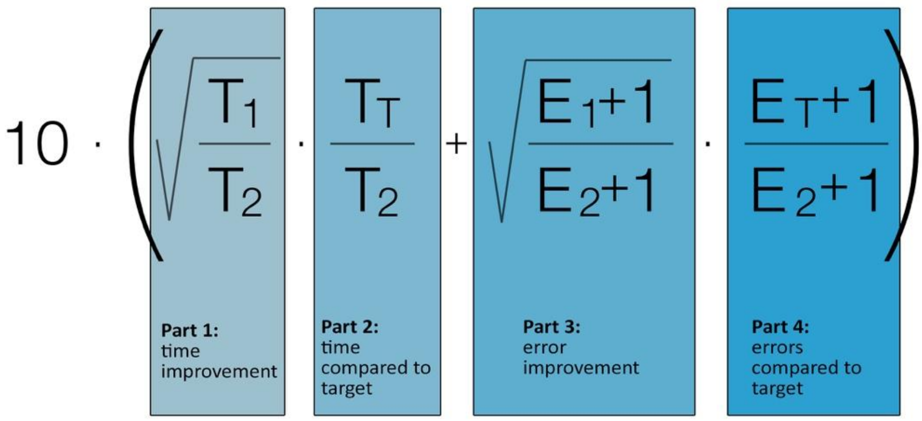
Considering this gap, we developed a novel algorithm that applies to dry-lab training and doesn't require any electronic data collection system. The Pi-score (performance improvement score) provides an objective measurement of performance improvement over a small number of trials, takes into account multiple variables and uses established quality standards as a benchmark. We performed our tests by applying it to tasks and rules from the European Basic Laparoscopic Urological Skills (EBLUS) [7], a training curriculum that, just like the Fundamentals of Laparoscopic Surgery (FLS) [8] or the AUA BLUS (6), is standardized, low cost and easily replicable. E-BLUS includes 4 tasks pertaining the most relevant basic skills in laparoscopy and can be taught in a standardized manner [9]. Therefore, it applied perfectly to our study, allowing us to focus on performance improvement in straight connection with personal skills and putting aside many potential biases, such as tutor's teaching ability, standardization of instruments, tasks or goals. In this study we evaluated the efficacy of our PI-score (performance improvement score) algorithm, by demonstrating its ability to score performance improvement with a reliability that is equal or even superior than a group of expert tutors. We also described the pi-score meaning, how it can be used for educational purposes and its future educational research applicability.

Materials and methods

Study of performance improvement was performed on the four E-BLUS tasks: peg transfer, circle cutting, needle guidance and knot tying. Rules and task goals were applied in accordance with what has already described in literature (7). We decided to study performance-improvement over 5 attempts because this time gap is already enough to identify an improvement, while still being applicable to the majority of courses. This way we wanted to offer full replicability to whomever wanted to double-check our study results.

Pi-score algorithm

The Pi-score is based on a novel algorithm, designed to create an objective and fast correlation between the results obtained by a trainee, during a HoT session. In order to produce the final score, the algorithm considers the three most relevant variables of each E-BLUS task: time-to-complete the task, target time as described by the EBLUS exam rules and number of errors. As shown in Figure 1, the algorithm calculation can be divided into four parts: time improvement, correlation with target time (Tt), error improvement and correlation with target number of errors (Et) per trial. Tt value is the time cut-off already defined for each EBLUS task, while Et value is usually set to 0 (no errors), in accordance with the EBLUS official exam rules. The first part of the algorithm considers the percentage of time-improvement between the first (T1) and the fifth attempt (T2).



$$10 \cdot \left(\sqrt{\frac{T_1}{T_2}} \cdot \frac{T_T}{T_2} + \sqrt{\frac{E_1+1}{E_2+1}} \cdot \frac{E_T+1}{E_2+1} \right)$$

Figure 11: PI-score calculation

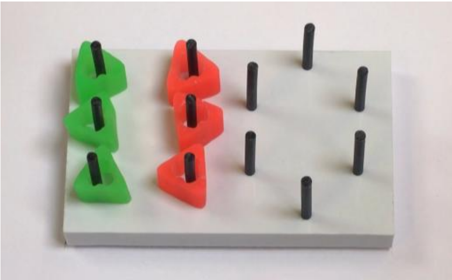
Thanks to this part of the algorithm, a higher difference between T2 and T1 will produce and increase of the final Pi-score. In the second part of the algorithm T2 value is put in comparison with the cut-off time of the task. Getting closer or performing below the cut-off time will rise up the PI-score. Rules and Cut-off times defined by EBLUS exam rules for each task are described in Figure 2. Part 3 of the calculation considers the percentage of improvement in relation to number of errors performed from the first to the fifth attempt. Part 4 considers errors on second trial compared to the ideal number of errors. Ideal number of errors is considered to be 0. In order to simplify the readability of the score, 4 cut-offs were defined to highlight different levels of performance improvement. The score could therefore indicate that the participant achieved a low (1), good (2), excellent (3) or outstanding (4) improvement. Depending on the difficulty to gaining improvement, that is intrinsic to each training task, the cut-offs were slightly trimmed to improve the sensitivity of the algorithm. The cut-offs defined for each task are summarized in Table 13.

Improvement	Peg transfer	Circle cutting	Needle guidance	Knot tying
Low	0 – 23	0 – 21	0 – 22	0 – 24
Average	23 – 30	21 – 35	22 - 30	24 – 36
Excellent	30 – 40	35 - 45	30 - 45	36 - 66
Outstanding	<40	<45	<45	<66

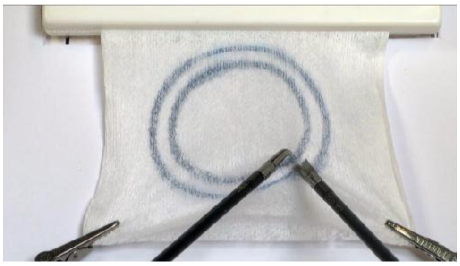
Table 13: Pi-score cut-offs

European-Basic Laparoscopic Urological Skills (E-BLUS) Exam

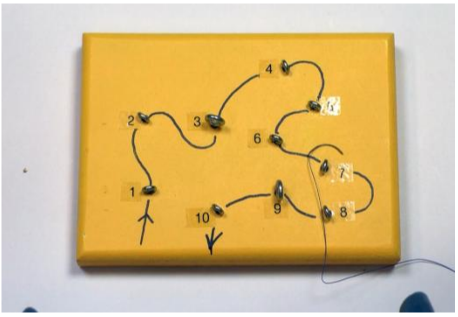
Task 1: Peg transfer (Target time 2:06 min:sec)

 2 graspers	Transfer all objects from one side to the other. Then opposite procedure. Start with non-dominant hand Transfer in mid-air	**Time start/stop:** First object is touched —until— Last object fallen **Error:** Number of dropped objects

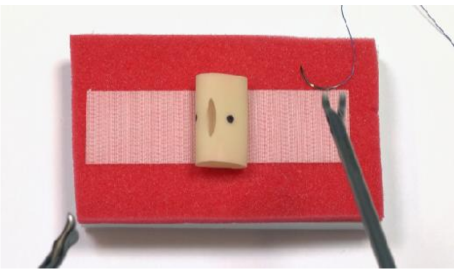
Task 2: Cutting a circle (Target time 2:31 min:sec)

 1 endoscopic scissor and 1 grasper	Cut the circle within the two lines. Start cutting within the circles.	**Time start/stop:** First cut —until— Circle comes loose **Error:** Cut beyond the lines Final product => enveloppe

Task 3: Needle guidance (Target time 4:28 min:sec)

 2 needle drivers, 1 suture with curved needle 50 cm	Guide the needle through the metal rings from 1 to 10 in the correct route. Do not straighten the needle by using too much force. When the needle is dropped, it is picked up again while time continues.	**Time start/stop**: Enter ring 1 —until— Needle passed ring 10 Errors: -

Task 4: Laparoscopic suture: single knot tying (Target time 6:00 min:sec)

 2 needle drivers (or 1 needle driver and 1 grasper), 1 endoscopic scissor and 1 suture with curved needle of 15cm	Place a single knot through the two dots. 3 throws on the suture: one double throw and two single throws ut suture 2-3 mm above the knot.	**Time start/stop:** Needle into rubber —until— Suture has been cut **Errors:** Stitch beyond 1 mm of mark No correct approximation Slipping knot Final product => enveloppe

Data collection

From 2014 to 2016 we collected data from 8 hands-on training courses. Each course lasted 2 days and was delivered by the same tutor (D.V.), using the same series of tips and tricks aimed to optimize performance on the EBLUS training tasks. Maximum 2 Participants were allowed in each training station, with time availability to complete 5 runs of each task. Run 1 and 5 were considered for this study. The first was performed as a baseline check with time and errors collection, before receiving any practical hint or guidance by the tutor. Afterwards, the tutor provided the participants with tips and tricks, to enhance and improve their performance. Time and errors from trial number 5 were considered to monitor performance improvement after tutor suggestions. A score-sheet [Table 2] was used to note the time taken for completion of tasks and the number of errors along the course. Participant demographics (age, gender), previous experience with laparoscopic and technological devices was also collected. The experience ranged from 1 to 5 (1- very low, 5-very high). For minimising bias and ensuring complete data collection the tutor personally oversaw this. By the end of each course, collected data were transcribed into the database in order to calculate the PI score of each participant.

Name Surname: ______________________ Age:____ PGY:_____

Date: __________

Laparoscopic expertise (1 poor – 5 very high): 1 2 3 4 5

General technology expertise (1 poor – 5 very high): 1 2 3 4 5

Have you ever taken part to a laparoscopic course? YES NO

Please use the template to take note of task time and number of errors during the course. The data will be used for research and quality purposes.

	Peg transfer Ref - 2:06	Circle cutting Ref – 2:31	Needle Guidance Ref – 4:28	Knot Tying Ref – 6:00
1st trial	min sec	min sec	min sec	min sec
N° errors	0 1 2 3 4 5 6	0 1 2 3 4 5 6	/	0 1 2 3 4 5 6
2nd trial	min sec	min sec	min sec	min sec
N° errors	0 1 2 3 4 5 6	0 1 2 3 4 5 6	/	0 1 2 3 4 5 6
3rd trial	min sec	min sec	min sec	min sec
N° errors	0 1 2 3 4 5 6	0 1 2 3 4 5 6	/	0 1 2 3 4 5 6
4th trial	min sec	min sec	min sec	min sec
N° errors	0 1 2 3 4 5 6	0 1 2 3 4 5 6	/	0 1 2 3 4 5 6
5th trial	min sec	min sec	min sec	min sec
N° errors	0 1 2 3 4 5 6	0 1 2 3 4 5 6	/	0 1 2 3 4 5 6

Table 14: score sheet.

To study the reliability of the score, a control arm of 14 experienced tutors was arranged. They were selected with at least four years of past experience as official EBLUS HoT proctor on behalf of the European Association of Urology (EAU). The tutors in the control arm received a database containing information (time + errors) from trials 1 and 5 for all the 50 participants enrolled. They were asked to subjectively score their perception of performance improvement for each of the 50 students, pertaining to the data provided. Their score had to be formulated on a scale from 1 to 4: 1-low improvement, 2-good improvement, 3-excellent improvement, 4-outstanding improvement. All collected information was then analysed for inter-rater reliability. The average scores from all tutors were then compared to the PI-score provided by our algorithm for each participant, in order to verify their statistical correlation.

Statistical Analysis

Data were summarized as mean and standard deviation (SD) or as absolute number, as appropriate. Cohen's Kappa statistics (for two raters and a binary response – i.e. with no degree of disagreement), Weighted Kappa statistics (which takes into account the degree of disagreement between two raters) or Kappa Statistics for multiple raters [10] were used to measure reliability among raters, as appropriate. One rater-removed analysis was applied to assess the influence by each single rater. The degree of the agreement analyzed by Kappa Statistics was scored according to the following scale: 0 – no agreement; 0-0.2 – slight agreement; 0.2- 0.4 – fair agreement; 0.4-0.6 – moderate agreement; 0.6-0.8 – substantial agreement; 0.8-1.0 – almost perfect agreement. A P-Value <0.05 was considered statistically significant. All statistical analyses were done by SPSS for Windows (version 22), IBM, Chicago, Illinois, USA or by STATA for Windows (version 13), Lakeway Drive, College Station, USA.

Results

50 participants were enrolled in the study from 8 courses. Mean age of the participants was 35.4 years (±8.7) with a male:female ratio of 43:7. Personal technological expertise was scored with an average of 3.7 (±1.1) out of 5, while personal laparoscopic expertise before the beginning of the course was scored with an average of 1.7 (±0.9) out of 5.

Regarding task 1 (Peg Transfer) agreement among the 14 expert raters about the performance improvement of the 50 participants, was "moderate" (Kappa=0.42) with Kappa ranging from 0.41 to 0.45 with one rater-removed analysis. On task 2 (Circle cutting), inter-rater agreement was "fair" (Kappa=0.27) with Kappa ranging from 0.26 to 0.29 with one rater-removed analysis. For task 3

(Needle guidance) expert agreement was "moderate" (Kappa=0.55) with Kappa ranging from 0.54 to 0.57 with one rater-removed analysis. On the last task (Knot tying) rater agreement was "fair" (Kappa=0.32) with Kappa ranging from 0.29 to 0.34 with one rater-removed analysis. As already pointed out, no relevant modifications of the results were found by one rater-removed analysis on the control arm.

The average per-participant scores collected from all the 14 expert tutors were then compared to the PI-scores produced by the algorithm, with the following results.

On task 1 the agreement between the expert average scores and the PI-scores was "almost perfect" (Kappa=0.85, 95% CI: 0.74-0.95). On task 2 the agreement between the expert average scores and the PI-scores was "moderate" (Kappa=0.46, 95% CI: 0.25-0.66). On task 3 the agreement between the expert average scores and the PI-scores was "almost perfect" (Kappa=0.92, 95% CI: 0.83-1.00). On task 4 the agreement between the expert average scores and the PI-scores was "substantial" (Kappa=0.65, 95% CI: 0.48-0.81). Agreement data are summarized in Figure 3.

The score produced by the algorithm during the study, formally the PI-score, had overall a minimum value of 7.3 and a maximum value of 102. No relevant correlation was found between the PI-score and the age of participants (P=0.159). No relevant correlation was found between the reported technological (P=0.843) or laparoscopic (P=0.417) background and the PI-scores obtained.

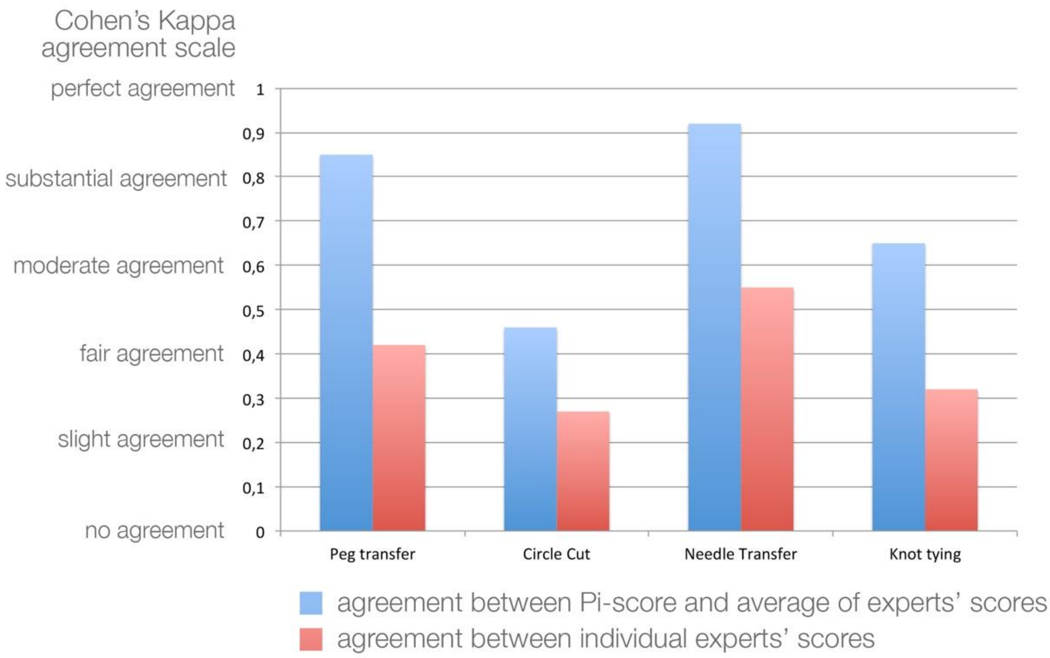


Figure 12: Agreement data summary

Discussion

The need to understand the learning process has guided educational goals for decades. The use of algorithms is now well established and has allowed the standardization of care and surgical approach in several medical areas [11,12]. As reported in literature, this improvement is most rapid in the beginning, with smaller improvements as the learning curve approaches a plateau [5,13]. For this reason, we found it valuable to test our algorithm on the very first series of training trials. The "proficiency-based training" suggested by Scott et al. [14] showed a complete success after a two-month period of free training for all the students enrolled but was not able to differentiate between those who had a faster acquisition of the technical skills and the others who didn't. Volpe et al. [15] explained how a 12-weeks training curriculum allowed to increase ability to perform the surgical steps of a prostatectomy (RARP) procedure, while in 2014 Abboudi et al. [16] systematically reviewed the relevance of learning curve in relation to the different urological procedures.

While performance improvement assessment is easier with the availability of several measurements over time, it becomes particularly difficult when the data available comes from a small number of attempts or from a restricted amount of training time. In order to study the feasibility of our algorithm, we had to compare it to what we consider the most reliable benchmark: experts' opinion. The lack of strict performance improvement evaluation parameters often leads anyway to an unstandardized result, as highlighted by our sample of experts. Indeed, the expert tutors, despite their huge assessment experience (> 4 years of HoT rating), provided highly variable results, which in the best scenario possible reached a "moderate" agreement (Kappa=0.42). This might be due to the lack of proper guidelines when it comes to performance improvement assessment. This result alone explains why it is so important today to have an automatic evaluation tool for performance improvement. Understanding and studying the outcomes of a training session is a core requirement when it comes to the optimization of an educational process. The introduction of such a tool can in fact allow evaluating not only the performance improvement of an individual, but also the course efficacy by simply calculating the average of participants' results. Moreover, the comparison between results by two groups with different proctors, simulators or curriculums could allow the analysis of a different variable (proctor's ability, simulator validity, curriculum efficacy), thus introducing several different uses of a simple scoring tool.

These reasons led us to the development of the PI-score algorithm. The PI-score analyses the overall performance improvement from one simulation trial to another one. This is differentiating it from a laparoscopic-skills assessment tool (as the E-BLUS curriculum itself) that is scoring participants' ability in an exact moment, but doesn't analyse their progression over time. For this reason, the PI-score is not introducing a novel way of assessing skills, but takes advantage of a well-established one, like the E-BLUS tasks with their rules, to score the efficacy of the entire session. An improvement of 50% does not always hold the same value, depending on how close the participant is to the task-specific time goal. For example, if we consider the peg transfer task (which is expected to be completed in less than 2.06 minutes with no errors) the PI-score will be 15.9 when the time is lowered down from 10 mins (trial 1) to 5 mins (trial 5) with no errors, while it will be 24.8 if the task time is brought down from 4mins (trial 1) to 2 mins (trial 5) with no errors. Pi-score will further be affected by errors along the trials

The use of the cut-offs has been applied to improve the readability of the results. While a simple number could be indeed confusing, we considered differentiating the results automatically, by applying different color-codes (Table 1) to each degree of performance improvement. Red was used for "low improvement", yellow for "good improvement", green for "excellent improvement", white for "outstanding improvement". This allowed structuring an easy-to-read summary page by the end of each course (example in Figure 4), which was well appreciated by the participants who had a fast and easy overview of their progression throughout the different training tasks. The use of the described cut-offs is anyway not only a visual hint, but also suggests how to further take advantage of the PI-score. A participant with repeated low PI-scores and high task times on the second measurement may, for example, demonstrate a low confidence with the given exercise and low chances to improve. On the opposite, low PI-scores coupled with times and errors that are already below the targets (eg. task 1: T1, 1:00 min; T2, 2:00 mins, no errors. PI-score: 17.4) may mean that the participant has already achieved the goals of the task and is therefore no more achieving any relevant improvement.

sex	PGY	Age	Peg transfer	Circle cutting	Needle guidance	Lap suturing
M	/	31	30,8	37,1	28,6	44,2
M	/	30	38,4	25,9	26,7	49,3
F	/	29	26,3	30,6	47,3	65,3
M	/	32	38,8	51,0	34,0	48,4
M	/	31	9,2	29,7	40,6	64,3
M	/	31	28,7	31,7	42,7	59,3
M	/	35	25,1	43,5	32,3	39,2
F	/	30	23,3	35,9	37,0	50,4

Figure 13: Pi-score summary page example

Limitations and future applications

Our study finds a clear limitation in the sample used: despite being enough for preliminary testing, more participants may be useful to consolidate cut-offs and reliability of the PI-score. The PI-score on the other side can be also seen as the first step of development for a machine learning system, which may select the most useful tips and tricks to provide to a student, based on his/her performance improvement scores. Pi-score might be also used to reinforce the assessment of a simulator or a curriculum as a pure teaching tool. To ensure full replicability of our study, Pi-score is available for free use on the webpage www.domenicoveneziano.it/piscore

Conclusions

In our study we demonstrated that evaluation of performance improvement is highly variable, even when it is formulated by a group of experts. We successfully developed an algorithm that is able to produce a score to objectively measure performance improvement, taking into account multiple variables over a limited number of samples. Our study demonstrated that the PI-score is feasible and equals the assessment reliability of a cohort of experts in relation to performance improvement after a small amount of training attempts.

References:

1) Parsons BA, Blencowe NS, Hollowood AD, Grant JR. Surgical training: the impact of changes in curriculum and experience. J Surg Educ 2011;68:44-51

2) Satava R. The future of surgical simulation and surgical robotics. Bulletin of the American College of Surgeons. 92. 13-9. 2007

3) S. A. Fraser, L. S. Feldman, D. Stanbridge, G. M. Fried
Characterizing the learning curve for a basic laparoscopic drill
Surg Endosc. 2005 Dec;19(12):1572-8. Epub 2005 Oct 17.

4) M. Macmillan, A. Cuschieri: Assessment of Innate Ability and Skills for Endoscopic Manipulations by the Advanced Dundee Endoscopic Psychomotor Tester: Predictive and Concurrent Validity – AM J SURG Vol. 177 March 1999

5) Feldman LS, Cao J, Andalib A, Fraser S, Fried GM. A method to characterize the learning curve for performance of a fundamental laparoscopic simulator task: Defining ''learning plateau'' and ''learning

rate''. Surgery, august 2009, vol 146, issue 2, pages 381-386
6) Kowalewski TM, Sweet R, Lendvay TS, et al. Validation of the AUA BLUS Tasks. J Urol. 2016 Apr;195(4 Pt 1):998-1005. doi: 10.1016/j.juro.2015.10.087. Epub 2015 Oct 23.
7) Brinkman WM, Tjiam IM, Schout BM, et al. Results of the European Basic Laparoscopic Urological Skills examination. Eur Urol. 2014 Feb;65(2):490-6. doi: 10.1016/j.eururo.2013.10.036. Epub 2013 Nov 6.
8) Peters JH, Fried GM, Swanstrom LL. Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. Surgery 2004;135:21-7. doi:10.1016/S0039-6060(03)00156-9
9) Somani BK, Van Cleynenbreugel B, Gozen A, et al.. The European urology residents education programme hands-on training format: 4 years of hands-on training improvements from the european school of urology. Eur Urol Focus 2018 Mar 14. pii: S2405-4569(18)30080-4.
10) Fleiss, J. L., B. Levin, and M. C. Paik. 2003. Statistical Methods for Rates and Proportions. 3rd ed. New York: Wiley
11) Hatib F, Jian Z, Buddi S, et al. Machine-learning Algorithm to Predict Hypotension Based on High-fidelity Arterial Pressure Waveform Analysis. Anesthesiology. 2018 Jun 12. doi: 10.1097/ALN.0000000000002300. [Epub ahead of print]
12) Ritzwoller DP, Hassett MJ, Uno H, et al. Development, Validation, and Dissemination of a Breast Cancer Recurrence Detection and Timing Informatics Algorithm. J Natl Cancer Inst. 2018 Mar 1;110(3):273-281.
13) Schmidt RA, LeeTD. Motor control and learning. A behavioural emphasis. 4th ed. Champaign (IL): Human Kinetics; 2005.
14) Scott DJ, Ritter EM, Tesfay ST, Pimentel EA, Nagji A, Fried GM. Certification pass rate of 100% for fundamentals of laparoscopic surgery skills after proficiency-based training - Surg Endosc (2008) 22:1887–1893
15) Volpe A, Ahmed K, Dasgoupta P, et al. Pilot Validation Study of the European Association of Urology Robotic Training Curriculum. Eur Urol. 2015 Aug; 68(2): 292-9. Doi: 10.1067j.eururo.2014.10.025. Epub 2014 Oct 31.
16) Abboudi H, Khan MS, Guru KA, et al. Learning curves for urological procedures: a systematic review. BJU Int 2014; 114: 617-629. Doi: 10.1111/bju.12315

CHAPTER 2.6

Veneziano D, Patruno G, Talso M, Tokas T, Proietti S, Porreca A, Kamphuis G, Biyani S, Emiliani E, Cepeda Delgado M, Maria de Mar Perez L, Miano R, Ferretti S, Macchione N q, Kallidonis P, Montanari E, Tripepi G, Ploumidis A, Lima E, Somani BK.

The Performance Improvement-score algorithm applied to EST s1 EAU protocol.

Data from ART in Flexible 2018.

The Performance Improvement-score algorithm applied to EST s1 EAU protocol. Data from ART in Flexible 2018.

Veneziano D [a,b,c], Patruno G[e], Talso M [f], Tokas T [g], Proietti S [h], Porreca A [i], Kamphuis G [j], Biyani S [k], Emiliani E [l], Cepeda Delgado M [m], Maria de Mar Perez L [n], Miano R [o], Ferretti S [p], Macchione N [q], Kallidonis P [r], Montanari E [s], Tripepi G [t], Ploumidis A [u], Lima E [a,b], Somani BK [v]

a) Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal
b) ICVS/3B's, PT Government Associate Laboratory, Braga/Guimarães, Portugal
c) Dept of Urology and Kidney Transplant, Grande Ospedale Metropolitano, Reggio Calabria, Italy

e) Dept of Urology, AO San Giovanni Addolorata, Rome, Italy
f) Dept of Urology, ASST Vimercate, Italy
g) Department of Urology and Andrology, General Hospital, Hall in Tirol, Austria.
h) Dept of Urology, San Raffaele-Turro Hospital, Milan, Italy
i) Dept of Urology, Policlinico Abano Terme, Italy
j) Dept of Urology, AMC University Hospital, Amsterdam, The Netherlands
k) Department of Urology, St. James's University Hospital Leeds Teaching Hospitals NHS, Leeds
l) Dept of Urology, Fundaciò Puigvert, Barcelona, Spain
m) Dept of Urology, Rio Hortega University, Valladolid, Spain
n) Dept of Urology, Centro de Cirugía de Mínima Invasión Jesús Usón, Caceres, Spain
o) Dept of Urology, Università Torvergata, Rome, Italy
p) Dept of Urology , AOU di Parma, Italy
q) Dept of Urology, Ospedale San Paolo, Milan, Italy
r) Dept of Urology, University of Patras, Greece
s) Dept of Urology, Policlinico Universitario, Milan, Italy
t) CNR IFC, U.O. of Reggio Calabria, Italy
u) Department of Urology, Athens Medical Centre, Athens, Greece
v) Department of Urology, University of Southampton, Southampton, UK
z) Dept of Urology, AO San Giovanni Addolorata, Rome, Italy

Word count: 2154

Abstract

Introduction and objectives

Pi-score has been proved to be reliable when it comes to measuring performance improvement during E-BLUS hands-on training sessions. Our study is aimed to adapt and test the score to EST s1 protocol, in consideration of its worldwide adoption for practical training.

Materials and methods

The Pi-score algorithm considers time measurement and number of errors from two different repetitions (first and fifth) of the same training task and compares them to the relative task goals, to produce an objective score. Data were obtained from the first edition of ART in Flexible, during 4 courses in Barcelona and Milan. Collected data were independently analysed by the experts for Pi assessment. Their scores were compared for inter-rater reliability. The average per-participant scores from all 14 proctors were then compared to Pi-score results. Kappa Statistics was used for comparison analysis.

Results

16 Hands-on Training expert tutors and 47 3rd year residents were involved. Concordance found between the 16 proctors' scores was the following: Task1=0.30 ("fair"); Task2=0.18 ("slight"); Task3=0.10 ("slight"); Task4=0.20, ("slight"). Concordance between Pi-score results and proctor average scores per participant was the following: Task1=0.74 ("substantial"); Task2=0.71 ("substantial"); Task3=0.46 ("moderate"); Task4=0.49 ("moderate").

Conclusion

Our study demonstrated that Pi-score can be adapted to EST s1, while still maintaining its calculation efficacy. Our algorithm successfully provided an objective score that equals

the average Pi assessment of a cohort of experts, in relation to a small amount of training attempts.

Introduction

Performing surgical training outside the operative room (OR) is considered fundamental in order to prepare for clinical practice[1]. The aim of lab and model training is focussed on surgical skills' details, improving technical performances to avoid using the patients as training platforms. Surgical training relies mostly on the curriculum, rather than on the simulator used[2] which means that the tutor plays a critical role in it.

Hands-on Training (HoT) has become increasingly popular in urological congresses and events and allows trainees and urologists to move their first steps under the guidance of tutors. To analyse their improvement, tutors observe the trainees along their trials and measure their achievements. This process is not objective or standardized and relies on personal expertise and opinion. Measuring performance improvement can be very useful to develop more reliable training methodologies and optimize any educational procedure. Laparoscopy has represented a giant leap in clinical practice, as well as surgical training, allowing clear view to the surgeon, its assistants and observers. Learning basic laparoscopic skills has been codified in renown and standardized curricula, such as E-BLUS and AUA-BLUS[3,4.]

Similarly, an endoscopic stone treatment training curriculum, named Endoscopic stone treatment step 1 (EST-s1) has been proposed to teach and assess trainees on endoscopic procedures[5,6]. This consists in 4 tasks, progressing from rigid cystoscopy to flexible ureteroscopy. In everyday practice, performance improvement of practical skills, both laparoscopic and endoscopic, is evaluated by expert's opinion, and is potentially biased by tutor's teaching ability, standardization of instruments, tasks or goals. The judgement of a group of experts, even when based on objective measures such has number of errors and time to complete tasks, can greatly vary[7].

In 2018 a new algorithm for performance improvement in laparoscopic skills has been proposed[7], that applies to dry-lab training and applies to any kind of data collection system, whether manual or digital. The Pi-score (performance improvement score) provides an objective measurement of performance improvement over a small number of trials, takes into account multiple variables and uses established quality standards as a benchmark. It has been tested on the EBLUS tasks and provided an objective score that equals the average Pi assessment of a cohort of experts.

In this study we evaluated the efficacy of the PI-score algorithm applied to the endoscopic steps of EST-s1, to evaluate its reliability against the average score of a group of expert tutors.

Materials and methods

Study of performance improvement was performed on the four EST-s1 tasks: rigid cystoscopy, flexible cystoscopy, semi-rigid UreteroRenoScopy (URS), flexible URS. Rules and task goals were applied in accordance with what has already described in literature[5]. The same criteria of 5 attempts of E- BLUS-PI was used, because this time gap is already enough to identify an improvement, while still being applicable to the majority of HoTs.

Pi-score algorithm

The Pi-score [5] has been designed to create an objective and fast correlation between the results of a HoT session. The Pi-score considers and measures the two most relevant variables of each task: number of errors and task completion time. Ideal timing and errors are defined following the established EST-s1 protocol[5]. Following the EBLUS Pi-score study, 4 cut-offs have been set to highlight different levels of performance improvement. A participant could therefore achieve a low (1), good (2), excellent (3) or outstanding (4) improvement. The cut-offs defined for each task are summarized in Table 15.

Improvement	Flex cyst	Rigid cyst	Semi-rigid URS	Flex URS
Low	0 – 25	0 – 30	0 – 31	0 – 31
Average	25 – 51	30 – 57	31 – 62	31 – 55
Excellent	51 – 85	57 – 90	62 – 95	55 – 81
Outstanding	<85	<90	<95	<81

Table 15: EST s1 Pi-score cut-offs

Data collection

In 2018 we collected data from 4 hands-on training courses, the ART in Flexible 1st edition, in Milan and Barcelona. Each course lasted 2 days and was delivered by the same tutor (D.V.), using the same series of tips and tricks aimed to optimize performance on the EST-s1 training tasks. The setup adopted during the courses included standard EST s1 simulators: the endoscopic simulator by Cook Medical (Bloomington, IN, USA) modified in accordance with the EST-s1 requirements, and the Coloplast K-Box (Humlebaek, DK). Participants allowed in each training station were maximum 2, with time availability enough to complete 5 runs of each task. The first trial was performed as a baseline check with time and errors' collection, before receiving any practical hint or guidance by the tutor. Afterwards, the tutor provided the participants with tips and tricks, to enhance and improve their performance. Run 1 and 5 were considered for this study. Time and errors from trial number 5 were collected to monitor performance improvement after tutor suggestions. This study was also considered to evaluate the efficacy of EST s1 as a teaching tool.

A score-sheet [Table 16] was used to note the time taken for completion of all tasks together with the number of errors along the course. Participants demographics (age, gender), previous experience with endoscopic procedures and technological devices was also collected with a Likert scale score ranged from 1 to 5 (1- very low, 5-very high). For minimising bias and ensuring complete data collection the tutor personally oversaw this phase.

Name Surname: ____________________ **Age:**____ **PGY:**_____
Date: _________

Endourology expertise (1 poor – 5 very high): 1 2 3 4 5
General technology expertise (1 poor – 5 very high): 1 2 3 4 5
Have you ever taken part to an endourology course? YES NO

Please use the template to take note of task time and number of errors during the course. The data will be used for research and quality purposes.

	Flex cystoscopy Ref – 0:36	Rigid Cystoscopy Ref – 0:41	Semi-Rigid URS Ref – 1:07	Flexible URS Ref – 3:43
1st trial	min sec	min sec	min sec	min sec
N° errors	0 1 2 3 4 5 6	0 1 2 3 4 5 6	/	0 1 2 3 4 5 6
2nd trial	min sec	min sec	min sec	min sec
N° errors	0 1 2 3 4 5 6	0 1 2 3 4 5 6	/	0 1 2 3 4 5 6
3rd trial	min sec	min sec	min sec	min sec
N° errors	0 1 2 3 4 5 6	0 1 2 3 4 5 6	/	0 1 2 3 4 5 6
4th trial	min sec	min sec	min sec	min sec
N° errors	0 1 2 3 4 5 6	0 1 2 3 4 5 6	/	0 1 2 3 4 5 6
5th trial	min sec	min sec	min sec	min sec
N° errors	0 1 2 3 4 5 6	0 1 2 3 4 5 6	/	0 1 2 3 4 5 6

Table 16: Scoring sheet used along the courses.

To study the reliability of the PI score, a control arm of 16 experienced tutors was arranged. Tutors with minimum 4 years of experience in Hands-on Training proctoring were selected. Each tutors was provided with a database containing information (time + errors) from trials 1 and 5 from each participant enrolled, in order to produce subjecting performance-improvement evaluation. Their evaluation had to be formulated on a scale from 1 to 4: 1-low improvement, 2-good improvement, 3-excellent improvement, 4-outstanding improvement. All collected information was then analysed for inter-rater reliability. The average scores from all tutors were then compared to the PI-score provided by our algorithm for each participant, in order to verify their statistical correlation.

Statistical Analysis

Data were summarized as mean and standard deviation (SD or as absolute number, as appropriate. Cohen's Kappa statistics (for two raters and a binary response – i.e. with no degree of disagreement), Weighted Kappa statistics (which considers the degree of disagreement between two raters) or Kappa Statistics for

multiple raters [10] were used to measure reliability among raters, as appropriate. One rater-removed analysis was applied to assess the influence by each single rater. The degree of the agreement analysed by Kappa Statistics was scored according to the following scale: 0 – no agreement; 0-0.2 – slight agreement; 0.2- 0.4 – fair agreement; 0.4-0.6 – moderate agreement; 0.6- 0.8 – substantial agreement; 0.8-1.0 – almost perfect agreement. A P-Value <0.05 was considered statistically significant. All statistical analyses were done by SPSS for Windows (version 22), IBM, Chicago, Illinois, USA or by STATA for Windows (version 13), Lakeway Drive, College Station, USA.

Results

47 PGY3 participants were enrolled in the study from 4 courses. Mean age of the participants was 29 years (± 1,2) with a male:female ratio of 32:15. Personal technological expertise was scored with an average of 2,8 (±0,9) out of 5, while personal endoscopic expertise before the beginning of the course was on average 2,5 (±0,8) out of 5.

On Flexible Cistoscopy, inter-rater agreement was "fair" (Kappa=0.30) with Kappa ranging from 0.21 to 0.35 with one rater-removed analysis. Regarding Rigid Cystoscopy performance improvement agreement among the 16 proctors was "slight" (Kappa=0.18) with Kappa ranging from 0.17 to

0.24 with one rater-removed analysis (P<0,005). For semi-flexible URS, expert agreement was "slight" (Kappa=0.10) with Kappa ranging from 0.09 to 0.11 with one rater-removed analysis (P<0,005). For the final task, Flexible URS, the agreement was "slight" (Kappa=0.19) with Kappa ranging from 0.11 to 0.23 with one rater-removed analysis (P<0,005). As already pointed out, no relevant modifications of the results were found by one rater-removed analysis on the control arm (P<0,005).

The average per-participant scores collected from all the 16 expert tutors were then compared to the PI-scores produced by the algorithm, with the following results.

On task 1 the agreement between the expert average scores and the PI-scores was "substantial" (Kappa=0.74, 95% CI: 0.58 - 0.89) (P<0,005). On task 2 the agreement between the expert average scores and the PI-scores was "substantial" (Kappa=0.71, 95% CI: 0.56 - 0.86) (P<0,005). On task 3 the agreement between the expert average scores and the PI-scores was "moderate" (Kappa=0.46, 95% CI: 0.27 – 0,65) (P<0,005). On task 4 the agreement between the expert average scores and the PI- scores was "moderate" (Kappa=0.49, 95% CI: 0,3 – 0,68) (P<0,005). Agreement data are summarized in Table 17.

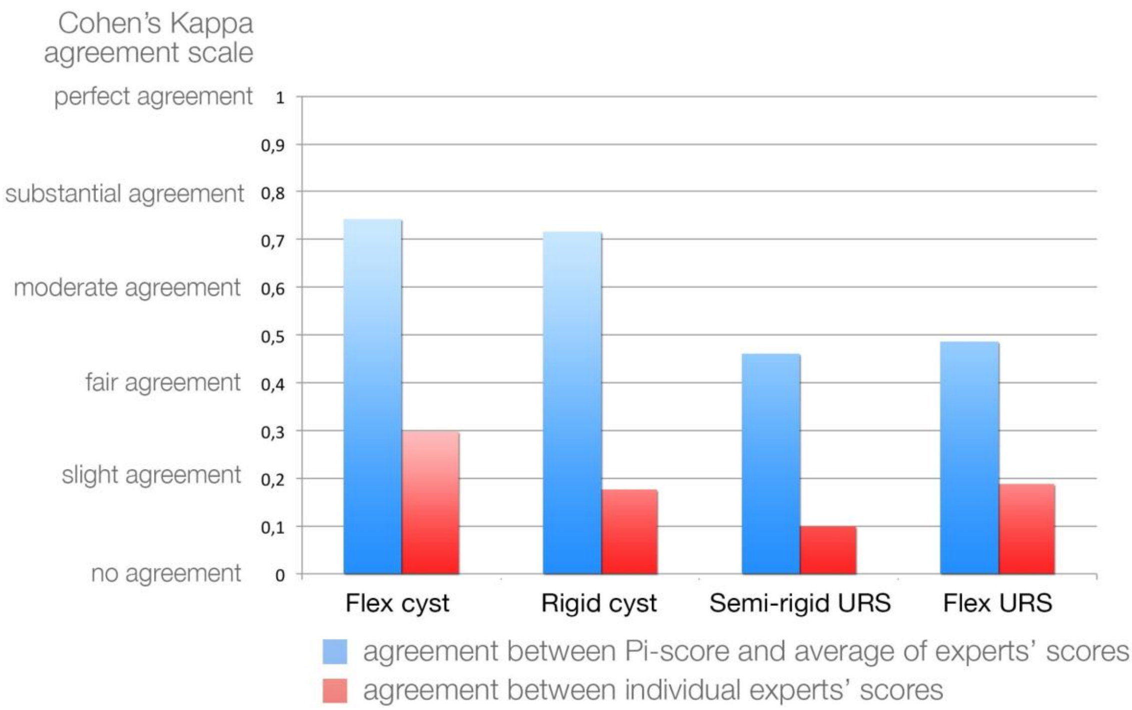


Table 17: summary of agreement data

Overall average Pi-score for all participants was 44,1 (±9,2). No relevant correlation was found between the PI-score and the age of participants. No relevant correlation was found between the reported technological or endoscopic background and the PI-scores obtained.

Discussion

Learning is a lifelong process. Training requires time and patience. The traditional principle established by Dr Halsted of "see one, do one teach one" surgical training method has crashed against a general reduction of the average hours/week, imposed by regulatory boards and governments[8]. Time of residency is limited by the increasing load of admin work, on calls and by the need to create competitive curriculum by means of publishing and research. This heavily impacts the hours dedicated to surgical training and the general satisfaction of the residents[9].

Training and getting ready to surgical procedures in a scientifically correct manner has become increasingly important, especially with the heavy legal implications that affect nowadays the healthcare environment. A curriculum-based approach, including simulation-based training, has proven to be effective in training surgeons to new procedures[10]. The Pi-score has been developed to objectively analyse performance improvement, optimize training methodologies and eventually spot those who demonstrate specific surgical talent. In the approach of a new skill, the improvement is higher in the beginning, while it becomes slower as the learning curve approaches a plateau [11]. For this reason, the original PI score was tested on the very first series of training trials [7]. The previous study already highlighted that expert opinion in this early learning phase is not reliable, with low agreement between the different raters especially because it's not usually feasible for a tutor to personally follow several trainees at the same time. This evidence was confirmed by the present study, which shows even lower agreement between the different tutors, regarding the performance improvement of the participants enrolled. This is probably due to the new protocol, EST s1, with new tips and tricks and learning curves, as opposed to the already renown E-BLUS [3,12]. In this case, the agreement between experts was poor, with the best result described as "fair" (Kappa=0.30) on task 1, rigid cystoscopy. The PI score was performing better and could predict experts' average scores with concordance measured as a "substantial" on task 1 and task 2 and "moderate" on task 3 and 4. For this reason the algorithm was used during the testing phase to select the participants who were demonstrating more progression and who were more suitable to proceed to advanced courses.

Moreover, analysing the Pi-score data from course to course allowed the faculty to improve and refine the teaching methodology and the hands-on training timing, which is core to the success of a training event. The overall average Pi-score registered between the 47 participants was 44,1 (±9,2), which corresponds to an "average performance improvement"(table 1) , with several pikes rated as "excellent improvement" (11/47 on task 1, 6/47 on task 2, 4/47 on task 3 and 17/47 on task 4), which added further validity evidence to EST s1 as a teaching tool.

Limitations and future applications

Our study confirms the findings of the previous one, but shares some of its limitations: the size of the sample used, which is intrinsic to the type of events selected. Allowing each participant to perform the same tasks several times requires time and simulation devices. Future studies could be conducted in training hospitals, equipped with the required simulation devices, allowing multi-centric data collection from more participants. Increasing data collection and tutor experience could also help to trim the quality cut-offs of the algorithm, making it more reliable for automated performance assessment.

The PI-score has proven its efficacy on different fields of urology and could be tested on different procedures and medical specialities or, in general, on any task requiring practice. Pi-score might be also used to reinforce the assessment of a simulator or a curriculum as a pure teaching tool. Pi-score is available for free use on the webpage www.domenicoveneziano.it/piscore

Conclusions

This study confirms our previous findings: subjective evaluation of performance improvement is extremely variable. The use of algorithms can help to objectively measure performance improvement. This study demonstrated that the PI-score, initially developed for laparoscopy, is a useful tool also to measure performance improvement during an EST s1 course, outperforming the judgement of a cohort of experts after a small amount of training attempts.

REFERENCES

1) Parsons, B. A., Blencowe, N. S., Hollowood, A. D. & Grant, J. R. Surgical training: The impact of changes in curriculum and experience. J. Surg. Educ. (2011). doi:10.1016/ j.jsurg.2010.08.004
2) Satava, R. The future of surgical simulation and surgical robotics. Bull. Am. Coll. Surg. **92**, 13–9

(2007).

3) Brinkman, W. M. et al. Results of the European Basic Laparoscopic Urological Skills Examination. Eur. Urol. **65**, 490–496 (2014).
4) Kowalewski, T. M. et al. Validation of the AUA BLUS Tasks. J. Urol. (2016). doi:10.1016/ j.juro.2015.10.087
5) Veneziano, D. et al. Validation of the endoscopic stone treatment step 1 (EST - s1): a novel EAU training and assessment tool for basic endoscopic stone treatment skills — a collaborative work by ESU , ESUT and EULIS. World J. Urol. **1**, (2019).
6) Veneziano, D. et al. Evolution and Uptake of the Endoscopic Stone Treatment Step 1 (EST-s1) Protocol: Establishment, Validation, and Assessment in a Collaboration by the European School of Urology and the
7) Uro-Technology and Urolithiasis Sections. Eur. Urol. **74**, (2018).
8) Veneziano, D. et al. Performance Improvement (Pi) score: an algorithm to score Pi objectively during E-BLUS hands-on training sessions. A European Association of Urology, Section of Uro- Technology (ESUT) project. BJU Int.(2019). doi:10.1111/bju.14621
9) Kotsis, S. V. & Chung, K. C. Application of the 'see one, do one, teach one' concept in surgical training. Plast. Reconstr. Surg. (2013).
10) doi:10.1097/PRS.0b013e318287a0b3
11) Carrion, D. et al. Current status of urology surgical training in Europe: an ESRU-ESU-ESUT collaborative study. World J Urol **Apr 13**, (2019).
12) Volpe, A. et al. Pilot Validation Study of the European Association of Urology Robotic Training Curriculum. Eur. Urol. (2015).
13) doi:10.1016/j.eururo.2014.10.025
14) Feldman, L. S., Cao, J., Andalib, A., Fraser, S. & Fried, G. M. A method to characterize the learning curve for performance of a fundamental laparoscopic simulator task: Defining 'learning plateau' and 'learning rate'. Surgery (2009). doi:10.1016/j.surg.2009.02.021
15) Somani, B. K. et al. The European Urology Residents Education Programme Hands-on Training Format: 4 Years of Hands-on Training Improvements from the European School of Urology. Eur.
16) Urol. Focus
17) (2018). doi:10.1016/j.euf.2018.03.002

CHAPTER 3

Discussion, conclusions and future perspectives

3. Discussion, conclusions and future perspectives

In the last decade surgical training has dramatically changed. Since 2010 what we call today "hands-on training" has become a worldwide accepted standard in Urology, with several studies describing its value. The first protocol adopted in Europe, the European Basic Laparoscopic Urological Skills, was heavily inspired by the American FLS, with slight adoptions to the urological field applied on behalf of the European Association of Urology (EAU). The present PhD project described the development, validation, uptake and refinement of the first ever hands-on training protocol "made-in-Europe": the Endoscopic Stone Treatment step 1. Its creation involved the EAU top experts in stone treatment, with the aim to produce a teaching methodology that could really be approached by anyone, as opposed to laparoscopy that is much more high-stakes oriented. The development process itself has been innovative, as the original "Full life-cycle curriculum development" template formulated by the simulation guru R.Satava, was actually used for the first time ever in this occasion. During the validation, Breakpoint-analysis was used for the first time to separate experts from novices in performance evaluation, allowing the identification of specific cut-offs. This detail, together with the adoption of the novel Messick's validation parameters, allowed the achievement of the "Best article of the Month" on the World Journal of Urology. The uptake of the EST s1 training and assessment protocol was incredibly fast, as it was strategically coupled to the already widespread E-BLUS, allowing the reach of 40 countries in just 2 years. Several companies started to adopt the protocol during their official courses, with simulators modified accordingly to meet the suggested requirements. The ART in flexible course was delivered in 2018 in Italy and Spain as a pilot project by EAU, using exclusively the EST s1, even to select the most talented residents to keep an eye on during the years to follow. The program was a success in 2018, with applications doubling the allowed number in 2019, just after one week from the opening of the registration website. Finally, the development of a dedicated Pi-score algorithm allowed to

check the actual reliability of the EST s1 as a training protocol, confirming that it is well structured and able to produce an average good improvement on participants, even just a limited series of training attempts.

Given what has been just described, the present PhD project might be considered the chronicle of a success, which exceeded any expectation of the developers involved and is today seen as a model to follow for future curriculum development. Richard Satava was right: "The simulator is just the tool, and it is the curriculum that will determine the training of the surgeon", because patients cannot continue to be the training platform.

EST s1 is not anyway the arrival, but the beginning of a reorganization in modern surgical training. The expertise collected along the last three years allowed us to understand the need for new training pathways, to guide residents along their early steps and up to clinical practice. The Endoscopic Stone Treatment curriculum will need to be integrated with the Step 2 and 3 as depicted by the Modular Hands-on Training template, to allow trainees learn how to perform parts of the procedure and then a full procedure before touching a real patient for the first time. Furthermore, more evidence on the drawback of training on real practice is needed, especially in order to create actual surgical certifications that might be recognized by scientific institutions, insurance companies and governments. This might bring to a completely new way of demonstrating surgical skills, with the final aim again to prevent or minimize complications. The introduction of Artificial Intelligence for performance assessment is another critical step in this evolution pathway, which will allow a further widespread of standardized training, with real time control of skill retention.

According to Moore's law, technology is been facing since the early nineties an exponential growth that is expected to still go on for decades. That's why it is easy to imagine that surgical training will become much more than an optional practice, but a true milestone and quality marker in the curriculum of any modern urologist.