Program for Laparoscopic Urologic Skills: A Newly Developed and Validated Educational Program

Irene M. Tjiam, Marjolein C. Persoon, Ad J. M. Hendrikx, Arno M. M. Muijtjens, J. Alfred Witjes, and Albert J. J. A. Scherpbier

OBJECTIVE To develop and evaluate a program for laparoscopic urologic skills (PLUS) to determine the face, content, and construct validation to achieve uniformity and standardization in training residents in urology.

METHODS The PLUS consists of 5 basic laparoscopic tasks. Three tasks were abstracted from the Fundamentals of Laparoscopic Surgery program, and 2 additional tasks were developed under continuous evaluation by expert urologists. Fifty participants were recruited from different hospitals and performed the final PLUS training. They all completed a questionnaire after performance. Three outcome parameters were measured: performance quality, time, and dropped objects. The relationship between laparoscopic experience and the outcome parameters was investigated.

RESULTS Of the 50 participants, 13 were students, 20 were residents, and 17 were urologists. Double-log linear regression analysis for all 5 tasks showed a significant effect (effect size range 0.53-0.82; $P < .0005$) for laparoscopic experience on performance time. Substantial correlations were found between experience and quality ratings (log-linear regression effect size 0.37; $P = .012$) and the number of dropped objects (Spearman correlation effect size 0.49; $P < .01$). The usefulness of the PLUS model as a training tool for basic laparoscopic skills was rated 4.55 on a scale from 1 (not useful) to 5 (useful) (standard deviation 0.58; range 3-5).

CONCLUSION The results of the present study indicated the face, content, and construct validity for the PLUS. The training is considered appropriate for use as a primary training tool for an entry test or as part of a step-wise training program in which basic and procedural laparoscopic skills are integrated. UROLOGY 79: 815–820, 2012.

With different forces eroding the feasibility of the conventional Halstedian model of "see one, do one, teach one," rapid changes are taking place in the teaching and learning of surgical skills. Ethical, legal, and financial considerations raise barriers to achieving the appropriate levels of mastery through training in the operating theater alone, thereby necessitating opportunities for new teaching and training methods. For quite some time, simulator-based training has been advocated as a logical and promising method for surgical skills training.[1-6]

As minimally invasive surgery has become more important in urologic practice, increased emphasis has been placed on laparoscopic education.[7-11] Laparoscopic surgery requires training in technical skills, such as hand-eye coordination and spatial awareness, which are different from the skills needed for open procedures.[3,12-17] Extensive studies have been undertaken to develop and evaluate simulators for minimally invasive techniques.[18] In the United States, the Fundamentals of Laparoscopic Surgery (FLS) program, based on the McGill Inanimate System for Training and Evaluation of Laparoscopic Skills, was introduced to teach fundamental laparoscopic knowledge and skills and has been extensively validated.[19-23] Dauster et al[24] reported evidence of its construct validity for urology training in the United States, where training and certification are incorporated into the residency program. In Europe, training methods for minimally invasive techniques outside the operating theater are being developed based on the FLS program and adapted to fit the European context. Because of these modifications, the adapted program must be validated before it can be implemented.

The aim of the present study was to examine the usefulness of a newly developed Program for Laparoscopic Urologic Skills (PLUS), which is based on the validated and widely used FLS program (face and content validity). In addition, we investigated whether the PLUS was able to distinguish between those with different experience

From the Catharina Hospital, Eindhoven, The Netherlands; Maastricht University, Maastricht, The Netherlands; and Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

Reprint requests: Irene Tjiam, M.D., Catharina Hospital Eindhoven, Michelangelolaan 2, 5623 EJ Eindhoven, The Netherlands. E-mail: irenetjiam@hotmail.com

Submitted: June 14, 2011, accepted (with revisions): January 10, 2012

0090-4295/12/$36.00
doi:10.1016/j.urology.2012.01.014

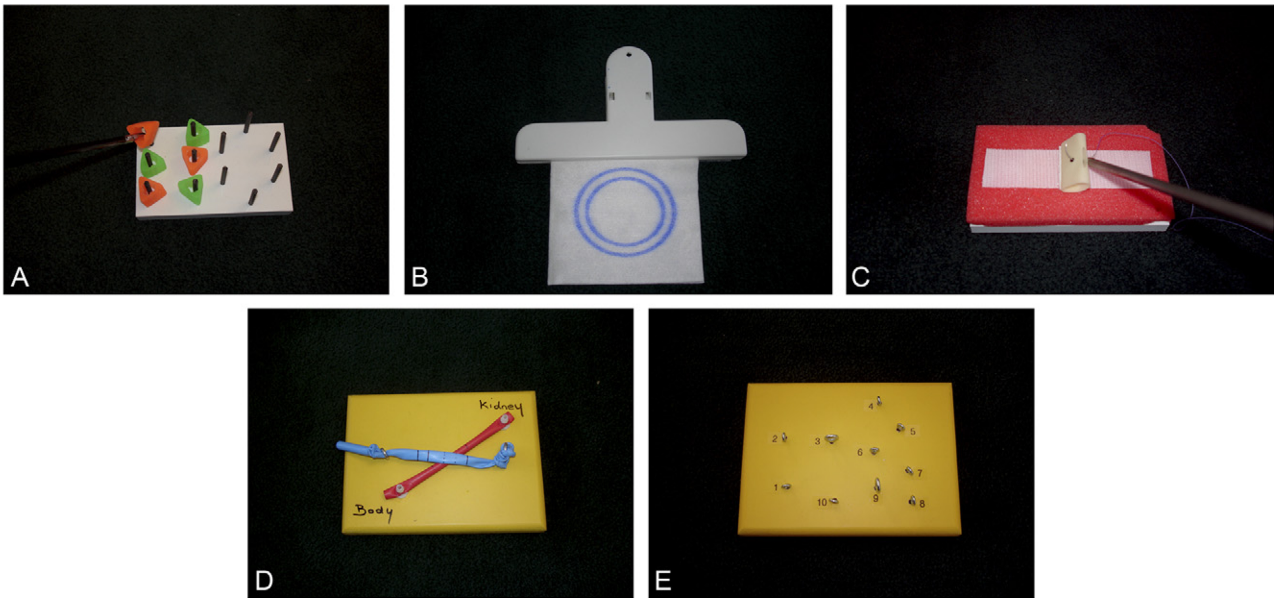

Figure 1. PLUS tasks. **(A)** Peg transfer, **(B)** cutting a circle, **(C)** single knot tying, **(D)** clip and cut, and **(E)** needle guidance.

levels in laparoscopy (construct validity). We sought evidence to support the hypothesis that the greater the level of laparoscopic experience, the shorter the time to complete a task, the greater the quality of task performance, and the fewer errors.

MATERIAL AND METHODS

Development of the PLUS

We submitted the original FLS training program to the judgment of 8 Dutch urologists, all from different hospitals, who are considered experts in the field of urologic laparoscopy, having performed >500 laparoscopic procedures and/or taught (inter-)national laparoscopy courses. After watching the theoretical modules on CD-ROM and performing the 5 original tasks on the box trainer, the experts expressed their opinions in a questionnaire considering the usefulness of each task for urologic practice.

Three tasks received a positive evaluation from all urologists (peg transfer, pattern cutting, and intracorporeal knot tying). However 5 (63.5%) and 7 (87.5%) of the 8 urologists considered the extracorporeal knot tying and the endoloop not relevant, respectively. Asked to suggest essential alternative tasks, the experts proposed a "needle guidance task" to train needle positioning and eye-hand coordination and a "clip-and-cut task" simulating clipping and cutting of the renal vessels during laparoscopic nephrectomy. After a thorough discussion, consensus on the exclusion of the 2 FLS tasks and inclusion of the 2 suggested tasks was reached in a meeting with the same experts. The 2 new tasks were then developed and continuously adjusted, in accordance with the suggestions and comments from these same experts, in cooperation with a skills laboratory technician of Catharina Hospital, Eindhoven. The final PLUS training includes 5 tasks of increasing complexity (Fig. 1).

The tasks were performed using the FLS box trainer, with a fixed-position video camera, 2 trocars with a fixed position, standard laparoscopic instruments (Karl Storz, Tuttlingen, Germany), a Hem-o-Lok clip applier (Teleflex Medical, Research Triangle Park, NC), and a monitor. For the knot-tying and needle guidance tasks, a suture of Polysorb 3-0 (Covidien, Dublin, Ireland) was used.

Validation

In the present study, we validated the PLUS according to the standardized steps of the validation process described by McDougal et al.[25] Face validity is defined as the "judgment of novices regarding the usefulness of the simulator," and content validity is defined as the "judgment of experts regarding the usefulness (appropriateness) of the simulator." Construct validity indicates whether the simulator is able to distinguish between experienced and inexperienced urologists.[25]

Participants

We recruited 50 participants from different hospitals and different levels of laparoscopic experience during laparoscopy courses and/or by electronic mail or by approaching them individually at their hospital. Included were urologists who perform laparoscopy, urology residents, and senior medical students (year 3 or later). Because it was logistically not feasible to transfer all participants to 1 hospital, the tests were held during the courses or in the participants' own hospitals.

All participants provided informed consent. No ethical approval was required because no patients were involved and the test results did not have a substantial effect on direct patient care.

Procedure

After explaining the tasks according to a standardized protocol, the participants were given 1 minute of practice time for each task to become familiar with the instruments and the task. For the actual test, the participants performed each task twice in succession (trials). All participants completed the PLUS tasks in the same order, from 1 to 5. Their performance was measured by recording the "time to complete the task," the number of "dropped objects" (task 1), and the "quality" of task performance (for tasks 2-4). The participants received no further guidance during the test.

For the first 2 tasks, the FLS protocol was used. Task 3 (single knot tying) was slightly little modified in the present study. We began timing when the needle was inserted into the rubber instead of the moment when the instruments were visible on the monitor. In the pilot study, it appeared that the students needed an excessive amount of time, up to 10 minutes, to position the needle in the needle driver before inserting it correctly in the rubber.

The following protocol was defined for the 2 newly developed tasks:

1. Clip-and-cut: this exercise is a simplified representation of the clipping and cutting of the renal vessels during nephrectomy. The trainee is required to place a loop with traction around a blue tube or "renal vein" to visualize the red tube or "renal artery." The trainee then places 3 clips on the artery with a Hem-o-Lok before making the cut. The same procedure is repeated for the vein. Timing began when the participant touched the loop and stopped when the artery was cut.
2. Needle guidance: the trainee was required to guide the needle through 10 metal rings following a set route. It was of no importance at which side the needle entered a ring. Timing began when the participant grasped the needle and stopped when the needle entered the last ring.

Performance was measured by recording the time with a stopwatch, the number of dropped objects in tasks 1, and the errors made in tasks 2-4. The time needed and dropped objects were recorded by a researcher. Two examiners, who were different from the researcher and unaware of the participants' names, gave their judgment on the errors independently. To judge the quality, we used a binominal 14-item checklist consisting of several relevant parameters from an existing and validated checklist for laparoscopic suturing.[26,27] For every error, a score of 0 was applied.

After the test, the participants were given a questionnaire concerning their demographics and baseline laparoscopic experience, defined as the "number of laparoscopic procedures performed independently or under supervision." The participants were also given a questionnaire asking them to give their opinion with regard to the usefulness of the 2 new tasks (clip-and-cut and needle guidance) and of the box trainer in general using a 5-point scale (1, not useful; 5, very useful).

Statistical Analysis

Because we expected the distribution of the variables "time to complete the task" and "experience" to be substantially skewed to the right, we performed linear regression analysis of the transformed versions: log(Time) and log(Experience). This transformation was not required for the variables "dropped objects" and "quality," because their distributions were not skewed. To avoid the problem that log(Experience) would be undefined when participants reported no experience, we used a slightly modified transformed value, log(Experience + 0.1). For this linear regression analysis, with a single dependent variable, the correlation coefficient could be used as an indicator of effect size (ES). The ES indicates to what extent an effect is of practical (clinical) importance.[28] To test the correlation between the number of dropped objects and experience, we used Spearman's correlation coefficient. The Statistical Package for the Social sciences, version 17 (SPSS, Chicago, IL) was used for all analyses. $P < .05$ was considered statistically significant.

Table 1. Opinions of all participants about PLUS usefulness and 2 newly developed tasks

Variable	Mean ± SD	Range
Clip-and-cut task		
Hand-eye coordination	4.33 ± 0.72	3-5
Instrument handling	4.53 ± 0.58	3-5
Needle guidance task		
Hand-eye coordination	4.73 ± 0.45	4-5
Instrument handling	4.66 ± 0.52	3-5
PLUS		
Hand-eye coordination	4.70 ± 0.50	3-5
Instrument handling	4.16 ± 0.80	2-5
3-Dimensional orientation	4.27 ± 0.86	2-5
Overall score on usefulness of simulator as educational tool	4.55 ± 0.58	3-5

PLUS, program for laparoscopic urologic skills.
Scale: 1, not useful; 5, very useful.

RESULTS

A total of 50 participants were included in the study: 13 students from the University of Maastricht, 20 urologic residents from 6 different hospitals, and 17 urologists from 10 different hospitals. The mean score for laparoscopic experience was 0 for students, 31 ± 18 for residents, and 439 ± 480 for urologists. Five participants were excluded from the analysis of construct validity, because they did not complete the whole PLUS. One participant was excluded from the analysis of face validity because he did not complete the final questionnaire.

Face and Content Validity

The participants' perceptions of the usefulness of the PLUS for training are listed in Table 1. Of all participants, 92% supported the inclusion of the clip-and-cut task, although 21% of the residents indicated that the loop element in this task offered no additional value. Also, 96% supported the inclusion of the needle guidance task. No significant differences were found among students, residents, and urologists with regard to the perceived usefulness of the PLUS trainer ($P = .137$, Kruskal-Wallis test).

Construct Validity

Substantial correlations were found between experience and time (double-log linear regression analysis, ES = 0.55-0.82; $P < .0005$; Fig. 2). The regression coefficient, correlation coefficient (ES), and R^2 per task are listed in Table 2. The decrease in time for experts versus novices (experience = 1000 vs experience = 0.1) ranged from 41% for tasks 1 and 2, 43% for task 5 (needle guidance), 46% for task 4 (clip-and-cut) to 74% for task 3 (single knot tying). For tasks 1, 4, and 5, consistent learning effects were found, with a decrease in time from the first to the second trial of 13% (task 1), 21% (task 4), and

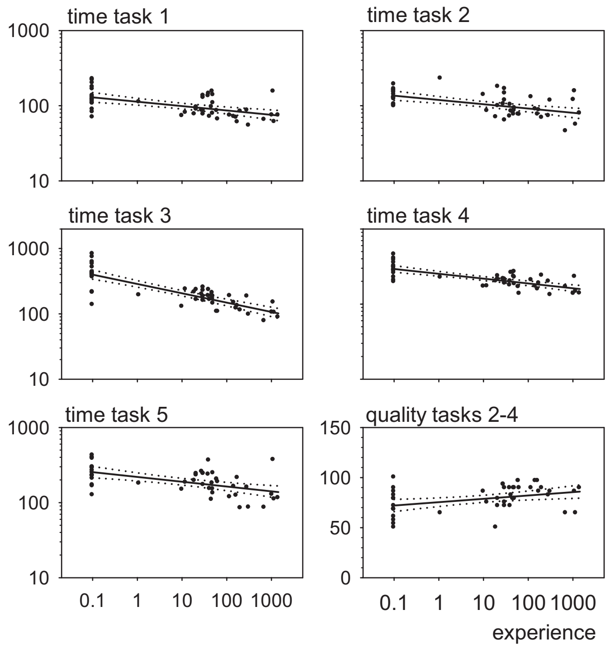

Figure 2. Linear regression analysis for "performance time versus experience" and "quality versus experience." Time task 1, peg transfer; time task 2, cutting gauze; time task 3, single knot tying; time task 4, clipping and cutting; and time task 5, needle guidance. Logarithmic scale used for time and experience; 95% confidential intervals indicated by dotted lines and each thick dot indicates a participant.

11% (task 5). For tasks 2 and 3, similar effects were found for novices (22% and 31% decrease, respectively), but not for experts, whose time to completion showed an increase of 8% for task 2 and a decrease of 1% for task 3.

A significant correlation was also found between experience and the quality of performance ratings (single-log linear regression analysis, ES = 0.37; $P = .012$). For the quality ratings, a relative increase of 18% was found for experts versus novices (Fig. 2).

Finally, a significant correlation was found between the level of experience and the number of objects dropped during task 1 (Spearman correlation coefficient, ES = 0.49; $P < .01$), where participants with more experience made fewer errors in terms of dropped objects.

COMMENT

The results of the present study have confirmed the face, content, and construct validity of a newly developed educational program, the PLUS, consisting of 5 basic tasks. Three tasks were abstracted from the FLS and 2 tasks were newly developed to fit the requirements of urology training programs.

Face and Content Validity

For establishing face and content validity, we based our cutoff point from previous studies by Sweet et al[29] and Schout et al.[30] A cutoff point of 3 on a 5-point scale (0, not useful; 5, very useful) was used to determine the acceptability of the PLUS trainer. All ratings were greater than 3.5 and those for the needle guidance task were consistently >4.5, suggesting that the PLUS can be considered a useful training method for laparoscopic urologic skills with acceptable face and content validity.

The clip-and-cut task was introduced to teach how to clip and cut the renal vessels. In particular for junior residents, who are not familiar with the Hem-o-Lok, it is considered of vital importance to learn the basics of this procedure. However, 21% of the residents thought the looping element should not be included in the task, because the traction given on the plastic tube was considered too hard and not realistic compared with the actual live traction with a loop on a renal vessel. However, all the urologists were in favor of inclusion of the looping element. This discordance of opinions suggests that this element should be evaluated carefully and possibly improved when the PLUS trainer is refined further. For example, using a longer loop would obviate the need for hard traction.

An important suggestion for additional improvement was routing directions for the needle guidance task, preferably guided through the rings by the left and right instrument alternately. We will take this into consideration at further refinement of the PLUS tasks.

The overall rating of the PLUS trainer was >4, with 3-dimensional orientation the only aspect to receive lower ratings owing to the fixed camera position. Finally, some participants remarked that the model was rather basic, because once the threshold performance criteria have been attained, no further improvement by the participant is possible.

Construct Validity

We used a different approach for establishing construct validity than conventionally to define the difference between the novice and expert. We did so, because the generally used broad qualification of expert or nonexpert in surgical skills and knowledge seems rather arbitrary and lacks the precision required for research. To avoid arbitrariness, we used the specific, continuous variable "experience," expressed as the number of laparoscopic operations performed and correlated this with the performance of basic laparoscopic tasks.

The strong correlation between "experience" and "time to complete the task" ($P < .0005$) and "quality" ($P = .012$) for all 5 tasks indicates overall better performance for more experienced participants, a finding that confirms our hypothesis and supports the construct validity of the PLUS trainer.

However, experience was not the only predictor of performance. Only 30% of the variance in outcome with regard to time for task 1 (peg transfer) could be explained by experience, leaving 70% to be explained by other variables, such as distraction, stress from time pressure, knowledge of how to handle the instruments correctly, and so forth. However, for single knot tying and the

Table 2. Linear regression analysis of "time versus experience" and "quality versus experience"

Time	Regression Coefficient (b)		Correlation Coefficient (Effect Size)	R^2
	b	Relative Decrease in Time From Novice to Expert (%)		
T1	−0.058*	42	−0.55*	0.30
T2	−0.057*	41	−0.55*	0.30
T3	−0.143*	73	−0.82*	0.68
T4	−0.064*	45	−0.73*	0.54
T5	−0.064*	45	−0.53*	0.28
PLUS quality	3.29†	18	0.37†	0.14

T, task.
* $P < .0005$.
† $P = .012$.

clip-and-cut task, experience explained 76% and 50% of the time results, respectively. This suggests that experience is a better predictor of performance with regard to the more applied or advanced skills, which seems logical, because more advanced skills offer more scope for improvement by training.

Study Limitations and Future Research

One limitation might relate to the definition of experience we used. We are aware of the probability that the total number of laparoscopic procedures (surgical and urologic) performed under supervision or independently was not accurately estimated. However, it can be assumed that the numbers given by the residents were accurate, because residents must record the number and specifications of the performed procedures in their portfolio. For the urologists, the number was estimated, especially for those who had performed >500 procedures.

Second, the students' opinions on the "usefulness of the PLUS trainer" should be interpreted with caution, because they had no relevant experience in laparoscopy; thus, their ratings might not be representative. In general, we have taken into account a certain amount of volunteer bias. All participants were willingly to perform the PLUS, which could imply that they already had a more positive attitude toward the PLUS than others who find laparoscopy less interesting.

Third, the relative increase of quality from novice to expert was only 18%, which might have been because we assessed errors after task performance (assessment of the materials) and not during the actual performance. However, the inter-rater reliability of this tool should still be established.

Finally, the results also suggest a substantial learning effect, especially for trainees with limited laparoscopic experience, although the learning effect seemed to vanish for highly experienced participants. However, no firm conclusions can be drawn with regard to individual learning curves on the basis of only 2 trials.

It would be worthwhile to focus future research on the learning curve within an individual and to identify the best training interval for the most efficient learning curve. The research protocol should not only focus on the amount of training, it should also consider aspects that might affect the efficiency of learning, such as different training intervals and the amount of mentoring. Furthermore, criterion validity of the PLUS trainer remains to be established, and a study should be conducted to determine the predictive value of PLUS performance for the performance of real-time laparoscopy. However, it is not recommended to transfer these basic skills directly to patients but to use the animal laboratory as an alternative. Currently, using procedural simulators as the reference standard is not optimal, because none of the simulators of that type have yet shown criterion validity for urologic laparoscopic procedures.

We recommend studying the next steps in training laparoscopic skills, using a variety of simulators, such as virtual reality simulators and live animal or cadaver models, to develop an educational laparoscopy curriculum. In addition, the reliability of the test should be researched and standards set to define a threshold of acceptable performance for use in the assessment and maintenance of basic laparoscopic skills.

CONCLUSIONS

In the present study we showed evidence for face, content, and construct validity of PLUS training to learn basic laparoscopic urologic skills. The PLUS shows promise as a primary training tool for a step-wise training method in which different aspects of basic and procedural skills are integrated. Additional studies should investigate the next steps for laparoscopic training. Also, additional research is needed to define the standards of acceptable performance that can be used for assessment or maintenance of basic laparoscopic skills.

Acknowledgment. To the 8 experts from Erasmus Medical Center, Radboud University Nijmegen Medical Centre, Canisius-Wilhelmina Hospital, Jeroen Bosch Hospital, Scheperhospital, St-Franciscus Gasthuis, St-Jansdal Hospital, and Rijnstate Hospital, who participated in the development of the PLUS trainer; to all the medical students from Maastricht University and the residents in urology and urologists who were willing to participate in this validation study; and to Mereke Gorsira for her editorial assistance, with thanks to all.

References

1. Autorino R, Haber GP, Stein RJ, et al. Laparoscopic training in urology: critical analysis of current evidence. *J Endourol*. 2010;24: 1377-1390.

2. Bashankaev B, Baido S, Wexner SD. Review of available methods of simulation training to facilitate surgical education. *Surg Endosc*. 2011;25:28-35.
3. Dunkin B, Adrales GL, Apelgren K, et al. Surgical simulation: a current review. *Surg Endosc*. 2007;21:357-366.
4. Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Engl J Med*. 2006;355:2664-2669.
5. Schout BM, Hendrikx AJ, Scheele F, et al. Validation and implementation of surgical simulators: a critical review of present, past, and future. *Surg Endosc*. 2010;24:536-546.
6. Scott DJ, Dunnington GL. The new ACS/APDS skills curriculum: moving the learning curve out of the operating room. *J Gastrointest Surg*. 2008;12:213-221.
7. Al-Shaiji TF, Kanaroglou N, Thom A, et al. A cost-analysis comparison of laparoscopic radical prostatectomy versus open radical prostatectomy: the McMaster Institute of Urology experience. *Can Urol Assoc J*. 2010;4:237-241.
8. Coelho RF, Rocco B, Patel MB, et al. Retropubic, laparoscopic, and robot-assisted radical prostatectomy: a critical review of outcomes reported by high-volume centers. *J Endourol*. 2010;24:2003-2015.
9. Eskicorapci SY, Teber D, Schulze M, et al. Laparoscopic radical nephrectomy: the new gold standard surgical treatment for localized renal cell carcinoma. *Sci World J*. 2007;7:825-836.
10. Ficarra V, Novara G, Artibani W, et al. Retropubic, laparoscopic, and robot-assisted radical prostatectomy: a systematic review and cumulative analysis of comparative studies. *Eur Urol*. 2009;55: 1037-1063.
11. Hemal AK, Kumar A, Kumar R, et al. Laparoscopic versus open radical nephrectomy for large renal tumors: a long-term prospective comparison. *J Urol*. 2007;177:862-866.
12. Fowler DL. Enabling, implementing, and validating training methods in laparoscopic surgery. *World J Surg*. 2010;34:621-624.
13. Laguna MP, de Reijke TM, de la Rosette JJ. How far will simulators be involved into training? *Curr Urol Rep*. 2009;10:97-105.
14. Melvin WS, Johnson JA, Ellison EC. Laparoscopic skills enhancement. *Am J Surg*. 1996;172:377-379.
15. Rassweiler J, Klein J, Teber D, et al. Mechanical simulators for training for laparoscopic surgery in urology. *J Endourol*. 2007;21: 252-262.
16. Rosser JC, Rosser LE, Savalgi RS. Skill acquisition and assessment for laparoscopic surgery. *Arch Surg*. 1997;132:200-204.
17. Wignall GR, Denstedt JD, Preminger GM, et al. Surgical simulation: a urological perspective. *J Urol*. 2008;179:1690-1699.
18. Kneebone R. Simulation, safety and surgery. *Qual Saf Health Care*. 2010;19(suppl 3):i47-i52.
19. Fraser SA, Klassen DR, Feldman LS, et al. Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system. *Surg Endosc*. 2003;17:964-967.
20. Fried GM, Feldman LS, Vassiliou MC, et al. Proving the value of simulation in laparoscopic surgery. *Ann Surg*. 2004;240:518-525.
21. McCluney AL, Vassiliou MC, Kaneva PA, et al. FLS simulator performance predicts intraoperative laparoscopic skill. *Surg Endosc*. 2007;21:1991-1995.
22. Peters JH, Fried GM, Swanstrom LL, et al. Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. *Surgery*. 2004;135:21-27.
23. Peters JH, Fried GM, Swanstrom LL, et al. Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. *Surgery*. 2004;135:21-27.
24. Dauster B, Steinberg AP, Vassiliou MC, et al. Validity of the MISTELS simulator for laparoscopy training in urology. *J Endourol*. 2005;19:541-545.
25. McDougall EM. Validation of surgical simulators. *J Endourol*. 2007; 21:244-247.
26. Aggarwal R, Hance J, Undre S, et al. Training junior operative residents in laparoscopic suturing skills is feasible and efficacious. *Surgery*. 2006;139:729-734.
27. Moorthy K, Munz Y, Dosis A, et al. Bimodal assessment of laparoscopic suturing skills: construct and concurrent validity. *Surg Endosc*. 2004;18:1608-1612.
28. Hojat M, Xu G. A visitor's guide to effect size. *Adv Health Sci Educ*. 2004;9:241-249.
29. Sweet R, Kowalewski T, Oppenheimer P, et al. Face, content and construct validity of the University of Washington virtual reality transurethral prostate resection trainer. *J Urol*. 2004;172(5 Pt. 1):1953-1957.
30. Schout BM, Bemelmans BL, Martens EJ, et al. How useful and realistic is the uro trainer for training transurethral prostate and bladder tumor resection procedures? *J Urol*. 2009;181:1297-1303.