JOURNAL OF ENDOUROLOGY
Volume 24, Number 11, November 2010

Pp. 1839–1843
DOI: 10.1089/end.2009.0612

Face and Content Validity of Transurethral Resection of Prostate on Uro Trainer: Is the Simulation Training Useful?

Shashikant Mishra, M.D., D.N.B., Abraham Kurien, M.D., D.N.B., Arvind Ganpule, M.D., D.N.B., Muthu Veeramani, M.D., M.Ch., R.B. Sabnis, M.D., M.Ch., and Mahesh Desai, M.D., FRCS (Eng), FRCS (Edin)

Abstract

Background: Uro Trainer (UT; Karl Storz GmbH, Tuttlingen, Germany), a virtual reality simulator for transurethral resection of prostate (TURP), has been infrequently validated. To ascertain the utility of such a trainer, we performed a basic face and content validity study.

Materials and Methods: Ten experts and nine novices (done more than 50 and less than 3 TURPs, respectively) performed a TURP on UT and rated simulator usefulness (seven items), realism (five items), and overall score (one item) on a Likert's 10-point scale. Scores of <6.0, 6.0 to 8.0, and >8.0 on the Likert scale 1 to 10 were considered not, slightly, and highly acceptable, respectively.

Results: Novices rated UT as more helpful than experts in the following aspects of face and content validity: usefulness general ($p = 0.0001$, statistically significant), hand–eye coordination ($p = 0.04$, statistically significant), material knowledge and skills ($p = 0.02$, statistically significant), spatial skills ($p = 0.003$, statistically significant), cystoscopy ($p = 0.002$, statistically significant), TURP (0.002, statistically significant), visual aspects ($p = 0.003$, statistically significant), and overall score ($p = 0.007$, statistically significant). One item of usefulness (coagulation) and three items of realism (tissue feel, depth feel, and capsule identification) failed to impress both novice and experts. UT usefulness domain was highly acceptable for 77.7% and slightly acceptable for 100% of the novices and experts, respectively. The general realism domain was highly and slightly acceptable for 33.3% and 66.6% of the novices while slightly acceptable for 100% of the experts.

Conclusion: Novice group found UT more useful than the experts group. Further modification is advisable to increase the realism of the UT.

Introduction

Surgical simulation may benefit residents to master transurethral resection of prostate (TURP). Until recently, these techniques were learned in an apprenticeship-training model by practicing on real patients.[1] The required skills can also be acquired in a risk-free environment by simulation training. To establish the efficacy of virtual reality (VR) training, rigorous validation studies need to be performed. There are, currently, two VR models for learning TURP. Face and content validity should be established before it is appropriate to proceed to higher levels of validity. The University of Washington TURP simulator has been described and validated extensively.[2–5] On the other hand, there is currently only a single published paper on Uro Trainer (UT), which failed to show the required basic preliminary validation.[6] The objective of the present study was to carry out a similar study for reassessment. The purpose of UT is to help novice understand the movements of the loop and hand–eye–foot coordination. In this way, the trainee's TUR completion rate would go higher without the supervising expert taking over. As per strict definition, novice should ideally meet the criteria of not doing any TURP and then learning the tricks of the procedure at the simulation model. In the present study, face validity is defined as the "judgment of novices regarding realism and usefulness of the simulator," whereas content validity is the "judgment of experts regarding realism and usefulness (appropriateness) of the simulator."[7]

Materials and Methods

The UT is a VR model that simulates TURP and transurethral resection of bladder tumor procedures. It is a plastic box linked to a computer hardware into which a resectoscope can be introduced. The development of the UT evolved out of a European community "ESPIRIT" research project for

Department of Urology, Muljibhai Patel Urological Hospital, Nadiad, Gujarat, India.

minimally invasive surgery simulation. It consists of a potent personal computer, a workstation with an integrated mechatronic unit (force feedback), and two monitors.[8] The central software generates and displays realistic images comparable, within diminutive tolerances, to those generated by endoscopic cameras. One monitor reflects the sight of the surgeon and the second demonstrates a sagittal, axial, or coronal view of the overall situation. Filling, empting, and flushing with irrigant are performed with genuine hardware incorporated into the instruments. Bleeding, responsive to fluid flow, is simulated accurately. The mechatronic unit provides realistic haptic sensations to the user. In addition to the prostate resection, it also incorporates bladder tumor resection procedures both under white and blue light endoscopy. Bare laser fibers can also be selected from the menu for therapeutic purposes (Fig. 1). TURP can be practiced with prostates of 34, 41, and 50 g.

A group of 19 participants involving 10 experts and 9 novices participated in the study. A pretask questionnaire, which provided demographic data, training status, and TURP-related questions, was collected from the participants. All the participants performed the resection procedure on the UT for a period of not less than 30 minutes or more till they felt comfortable to carry out the test procedure. All the participants then performed a TURP on UT with 41 g of prostate. The setting of 41 g was the intermediate value of the three sizes of prostate TURP module available on the UT. This was chosen for the study to bridge any differences, which may have been in smaller and larger glands TURP module. This was fixed for all the participants to bring uniformity of the procedure. The time limit set on the UT was 30 minutes for the test procedure; irrespective to whether or not resection was complete. Subsequent to performing the test procedure, they were allowed to fill up the questionnaire sheet before exiting the laboratory. In essence, there was hardly any time for elaborate thinking on individual domains of the realism and usefulness.

Haptic feedback was deleted from the task to avoid the test bias and to make comparison with the lone published study. Experts were defined as urologist having performed more than 50 TURPs. Novice were defined as trainees having done less than three endoscopic resection procedures including bladder tumor and prostate.

The participants were isolated for the procedure and filing of the questionnaire. No interindividual discussion was allowed to avoid the influence of information from others to result in correspondence bias for evaluating each domains. Two sets of pretask and posttask questionnaires were completed by the participants for analysis. The pretask questions consisted of an open-ended information regarding the relevance of the dry laboratory training. The posttask questions consisted of 13 items asking participants to rate UT usefulness (seven items), realism (five items), and overall score (one item) on ordinal 10-point scales (1—not at all useful/not at all realistic; 10—very useful/very realistic). Scores of <6.0, 6.0 to 8.0, and >8.0 on the Likert scale 1 to 10 were considered not acceptable, slightly acceptable, and highly acceptable, respectively.

Discrete variables were compared with the chi-square test and presented as numbers and percentages. The Student's t test was performed to evaluate relationship among overall questions. A p-value of <0.05 was considered to indicate statistical significance for all tests. All statistical analyses were conducted using SPSS® version 15.0.

FIG. 1. Uro Trainer (Karl Storz GmbH, Tuttlingen, Germany).

Results

A total of 19 participants completed the questionnaires. On the basis of the number of TURP procedures performed, 10 and 9 participants were classified as experts and novices, respectively. Demographics are presented in Table 1. The participants were from the western part of the country. Majority of the experts were practicing urologists and the trainees were residents in the first year.

Novices were younger than the experts ($p < 0.001$). Pretask open-ended questions are shown in Figure 2. All the participants, irrespective of their expert and novice status, agreed to have faced problems in acquiring necessary skills during the initial learning curve of the TURP in real OR situation. The

Table 1. Demographics of Expert and Novice

	Expert (n = *10*)	*Novice* (n = *9*)	p-*Value*
Age (years)	41.2 ± 8.6	29.5 ± 1.1	<0.001[a]
No. of transurethral resection of prostate	249.5 ± 204.8	1.4 ± 0.5	<0.001[a]
Faced problems in the initial learning curve	100%	100%	
Would prefer simulation training if available	80%	100%	
Willing to purchase the simulator for training	70%	100%	

[a]Statistically significant.

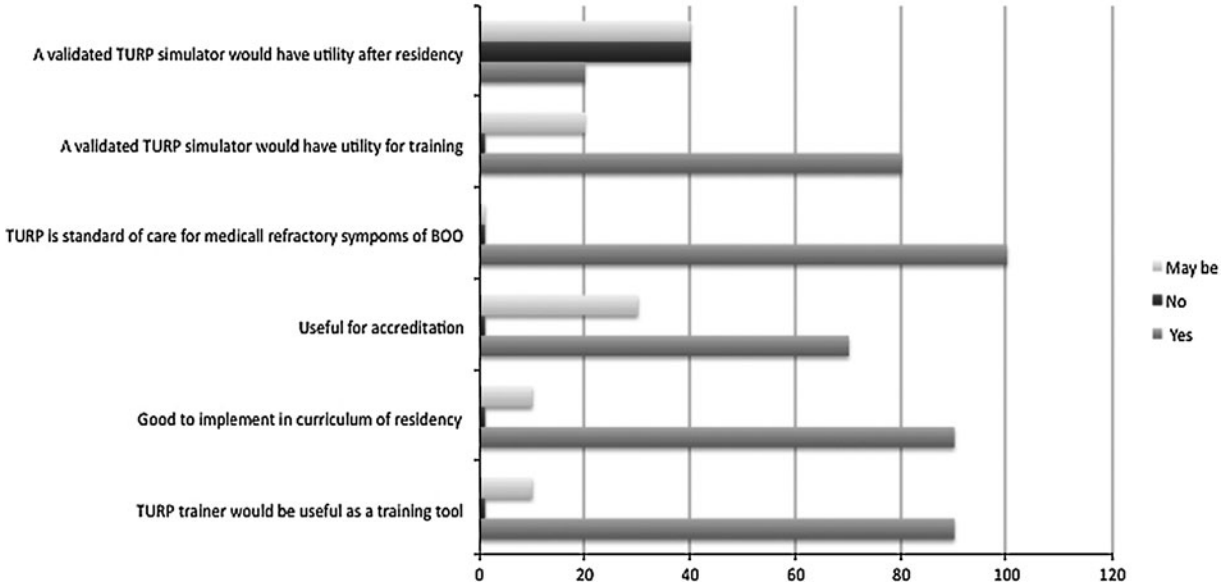

FIG. 2. Pretask open-ended questionnaire.

majority of the participants had doubt of a validated TUR model having utility after the training, but agreed that it still had utility during the training. We acknowledge that there may be a selection bias toward such a response because subjects who chose to use the simulator and participate in the study were probably more likely to have such an attitude.

Table 2 shows the mean ± standard deviation of the posttask questionnaire rating by the expert and novice. Novice when compared with experts differed in opinion in the domains of usefulness general, hand–eye coordination, material knowledge and skills, spatial skills, cystoscopy, TURP, visual aspects, and overall score. Four of five domains of realism were similar for both.

For novice, measured against six criteria of useful domain (Fig. 3), UT was highly acceptable for 77.7%, 66.6%, 44.4%, and 55.5% for the domains of usefulness general, hand–eye coordination, material knowledge and skills, and spatial skills, respectively. It was slightly acceptable for 88.8%, 77.7%, and 66.6% for domains of cystoscopy, coagulation, and TURP, respectively. In contrast, for experts, UT was slightly acceptable for 100%, 70%, 90%, 80%, 90%, and 80% for the domains of usefulness general, hand–eye coordination, material knowledge, spatial skills, coagulation, and TURP, respectively. It was not acceptable for 80% of experts for domains of cystoscopy.

For novice, measured against five criteria of realism, UT was not acceptable for 100% each for domains of tissue feel, depth feel, and capsule identification and was slightly acceptable for 66.6% and 55.6% of novice for realism in general and visual aspects. Similarly, UT was not acceptable for 100% of experts for each of domains of tissue feel, depth feel, and capsule identification and was slightly acceptable for 100% and 90% of experts for realism in general and visual aspects.

Discussion

Simulation offers an appealing adjunct for learning TURP skills.[9–11] Properly constructed simulation models may provide trainees with increased opportunities to learn TURP in a safe and effective manner.[6] Simulator's ability to train is authenticated when it can prove its worth by rigorous validation studies. In the present study, we established the basic face and content validity of the UT. Adequate face and content validity would justify further higher level of validation studies.

There are no clear-cut guidelines to objectively measure face and content validity. Studies in which face and content validity is measured quantitatively use rather arbitrary cutoff points.[6] In the present study, we used the scale as described by Schout et al,[6] to avoid any conclusion difference. In their study, the modified 5-point Likert scale used by Sweet et al[2] was converted to a 10-point scale. Scores of 6.0 and 8.0 on the

Table 2. Mean ± Standard Deviation Values of the Posttask Questionnaire Rating by the Expert and Novice

Serial no.	*Domain*	*Expert*	*Novice*	p-*Value*
1	Usefulness general	6.4 ± 0.52	8.4 ± 0.88	0.0001[a]
2	Hand–eye coordination	7.4 ± 0.69	8.0 ± 0.70	0.04[a]
3	Material knowledge and skills	6.7 ± 0.67	7.4 ± 1.00	0.02[a]
4	Spatial skills	6.1 ± 0.73	7.4 ± 1.00	0.003[a]
5	Cystoscopy	5.2 ± 0.63	6.2 ± 0.66	0.002[a]
6	Coagulating	6.4 ± 0.69	6.7 ± 0.66	NS
7	Transurethral resection of prostate	6.4 ± 0.69	7.3 ± 0.50	0.002[a]
8	Realism general	6.3 ± 0.48	6.8 ± 0.92	NS
9	Visual aspects	6.5 ± 0.70	7.4 ± 0.52	0.003[a]
10	Tissue feel	1.0 ± 0.66	0.2 ± 0.44	NS
11	Depth feel	0.5 ± 0.71	0.4 ± 0.52	NS
12	Capsule identification	0.4 ± 0.51	0.1 ± 0.33	NS
13	Overall score	6.5 ± 0.52	7.4 ± 0.52	0.007[a]

[a]Statistically significant.
NS = not significant.

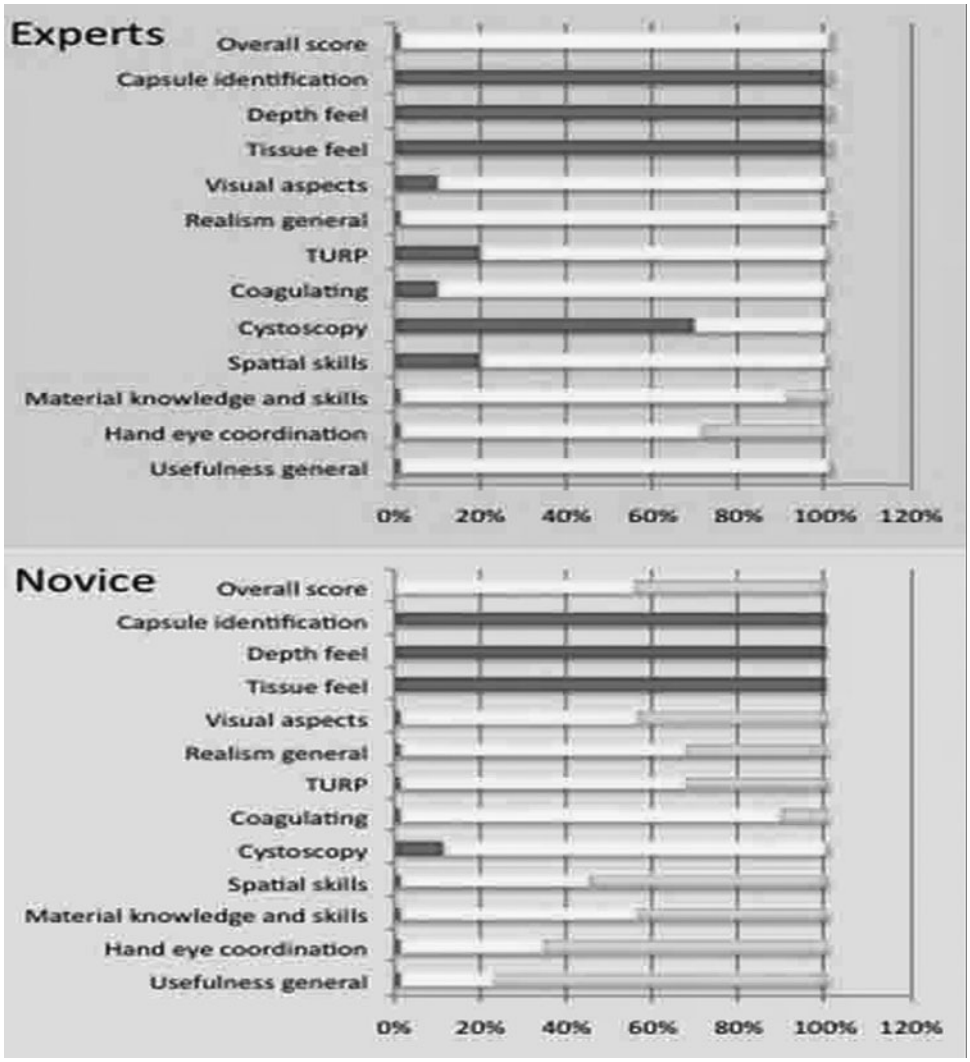

FIG. 3. Measures of validity and difference of opinion between expert and novice. Measured against Ayodeji criteria: average score of 8 or above (grey, highly acceptable), 6 to 8 (white, slightly acceptable), and 6 or below (black, not acceptable). TURP = transurethral resection of prostate.

converted scale were considered slightly acceptable and moderately acceptable, respectively. Ayodeji et al[12] considered a score of 8.0 or higher on a 10-point scale as a positive score.

In accreditation of training programs, novice scenario is met by the trainees who do not have any experience in the parameter to be studied. In TUR training, the independent TURP for the novice criteria to be met is zero. One of the secondary objectives of the present study was also to evaluate the difference in the response of the questionnaires by the novice *vis-à-vis* expert. This is because, the novice population is the target population. What may seem ridiculous to the expert may not necessarily be so for the novice. However, if the domain questionnaire result is similar for both the population, then the reliability of the domain in the model is high.

There was a marked difference of opinion between the two groups on certain aspects of usefulness. Novices perceived UT to be more useful in all domains except cystoscopy. Why was there a difference of opinion among the two in contrast to the study by Schout et al,[6] in which the difference was not statistically different. First, the novice group was more interested in learning the procedure. It is known that face validity is inferior to content validity. Novice may be unaware of the likely loop holes of the simulation because of the lack of exposure and so they rate the validity higher. Second, the participant group with respect to the published paper was different. Unlike the attendees of the world congress where time factor of the participant could jeopardize the desire to participate in the study and would negatively skew the results, we chose to include participants who voluntary took part in the study and had no time limitation. They had ample time to think about the various constructs to do TURP and identify the deficiency in a specific construct to be improved upon. Third, the novice definition was followed. Most authors suggest that a mean number of 40 to 50 procedures for each urologist are the minimum to achieve satisfactory skills on TURP.[13] UT is then bound to have inferior face validity results, which actually reflect the content validity. Hence, unlike the published study, we chose the novice group having performed less than three procedures.

Novice and expert view on realism for the UT was not significantly different. There was a 100% negative concordance with the realistic domains tissue depth, depth feel, and capsule identification. To our surprise, despite such a strong opinion against the majority of realism domains, the realism in general domain was slightly and highly acceptable for 100% and 0% versus 66.6% and 33.3% of experts and novice, respectively. One of the assumptions for this difference in opinion may be the overlapping of the strong usefulness domain diluting the negative aspect of realism. The second possible explanation may be by the fact that we did not ask specifically of the resectoscope and the cautery pad. If the realism of these aspects are highly positive, then the overall realism increases.

The main limitation of the study is a small sample size, which limited statistical power for some desired variant domains. We were limited by this disadvantage because it was conducted in a single center. The intent of short sample size was to test the hypothesis in a small number of subjects first. However, it still can be kept as a pilot project for further such studies to be done with adequate power sample size. The novice group, albeit small, who were more interested and enthusiastic for the response, could have positively skewed the results.

Conclusion

In our study, UT was found to have satisfactory usefulness domains but lacked the majority of realism domains. Hence, further modification in the realistic domains is required. On the basis of our study, manufacturers need to improve upon the constructs in domains such as tissue feel, depth feel, and capsule identification before undertaking any attempt at further validation.

Aknowledgments

The authors thank Dr. Pradeep Ganatra (coordinator of workshops, Jayramdas Patel Academic Centre, Nadiad) for arranging the participant sessions and Dr. Akhilesh (incharge, dry and wet lab) for helping in data collection.

Disclosure Statement

No competing financial interests exist.

References

1. Satava RM. Historical review of surgical simulation—A personal perspective. World J Surg 2008;32:141–148.
2. Sweet R, Porter J, Oppenheimer P, Hendrickson D, Gupta A, Weghorst S. Simulation of bleeding in endoscopic procedures using virtual reality. J Endourol 2002;16:451–455.
3. Sweet R, Kowalewski T, Oppenheimer P, Weghorst S, Satava R. 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.
4. Rashid HH, Kowalewski T, Oppenheimer P, Ooms A, Krieger JN, Sweet RM. The virtual reality transurethral prostatic resection trainer: Evaluation of discriminate validity. J Urol 2007;177:2283–2286.
5. Wignall GR, Denstedt JD, Preminger GM, et al. Surgical simulation: A urological perspective. J Urol 2008;179:1690–1699.
6. Schout BM, Bemelmans BL, Martens EJ, Scherpbier AJ, Hendrikx AJ. How useful and realistic is the Uro Trainer for training transurethral prostate and bladder tumor resection procedures? J Urol 2009;181:1297–1303.
7. McDougall EM. Validation of surgical simulators. J Endourol 2007;21:244–247.
8. Reich O, Noll M, Gratzke C, et al. Stief. High-level virtual reality simulator for endourologic procedures of lower urinary tract. Urology 2006;67:1144–1148.
9. Manyak MJ, Santangelo K, Hahn J, et al. Virtual reality surgical simulation for lower urinary tract endoscopy and procedures. J Endourol 2002;16:185–190.
10. Gomes MP, Barrett AR, Timoney AG, Davies BL. A computer-assisted training/monitoring system for TURP structure and design. IEEE Trans Inf Technol Biomed 1999;3:242–251.
11. Oppenheimer P, Gupta A, Weghorst S, Sweet R, Porter J. The representation of blood flow in endourologic surgical simulations. Stud Health Technol Inform 2001;81:365–371.
12. Ayodeji ID, Schijven M, Jakimowicz J, Greve JW. Face validation of the Simbionix LAP Mentor virtual reality training module and its applicability in the surgical curriculum. Surg Endosc 2007;21:1641–1649.
13. Mebust WK, Holtgreve HL, Cockett ATK, et al. Transurethral prostatectomy: Immediate and postoperative complications. A cooperative study of 17 participating institutions evaluating 3885 patients. J Urol 1989;141:243–247.

Address correspondence to:
Mahesh R. Desai, M.D., FRCS (Eng), FRCS (Edin)
Medical Director and Managing Trustee
Muljibhai Patel Urological Hospital
Dr. Virendra Desai Road
Nadiad, Gujarat 387001
India

E-mail: mrdesai@mpuh.org

Abbreviations Used

NS = not significant
TURP = transurethral resection of prostate
UT = Uro Trainer
VR = virtual reality