International Journal of Surgery 10 (2012) 163–166

Contents lists available at SciVerse ScienceDirect

International Journal of Surgery

journal homepage: www.theijs.com

Original research

Face validity, construct validity and training benefits of a virtual reality turp simulator

Elizabeth Bright [a,*], Samuel Vine [b], Mark R. Wilson [b], Rich S.W. Masters [c], John S. McGrath [a]

[a] Exeter Surgical Health Services Research Unit, Royal Devon & Exeter NHS Foundation Trust, Barrack Road, Exeter, Devon EX2 5DW, UK
[b] College of Life and Environmental Sciences, University of Exeter, St Luke's Campus, Heavitree Road, Exeter, Devon EX1 2LU, UK
[c] Institute of Human Performance, University of Hong Kong, Pokfulam, Hong Kong

A R T I C L E I N F O

Article history:
Received 24 September 2011
Received in revised form
27 November 2011
Accepted 13 February 2012
Available online 23 February 2012

Keywords:
TURP
Validity
Virtual reality
Simulation

A B S T R A C T

Objective: To assess face validity, construct validity and the training benefits of a virtual reality TURP simulator.
Method: 11 novices (no TURP experience) and 7 experts (>200 TURP's) completed a virtual reality median lobe prostate resection task on the TURPsim™ (Simbionix USA Corp., Cleveland, OH). Performance indicators (percentage of prostate resected (PR), percentage of capsular resection (CR) and time diathermy loop active without tissue contact (TAWC) were recorded via the TURPsim™ and compared between novices and experts to assess construct validity. Verbal comments provided by experts following task completion were used to assess face validity. Repeated attempts of the task by the novices were analysed to assess the training benefits of the TURPsim™.
Results: Experts resected a significantly greater percentage of prostate per minute ($p < 0.01$) and had significantly less active diathermy time without tissue contact ($p < 0.01$) than novices. After practice, novices were able to perform the simulation more effectively, with significant improvement in all measured parameters. Improvement in performance was noted in novices following repetitive training, as evidenced by improved TAWC scores that were not significantly different from the expert group ($p = 0.18$).
Conclusions: This study has established face and construct validity for the TURPsim™. The potential benefit in using this tool to train novices has also been demonstrated.

1. Introduction

Over the past decade, there have been significant changes in the UK specialist urological training programme. A reduction in the length of specialist training and the shortening of working hours, due to the implementation of the European Working Time Directive, has placed increasing demands on urology trainees to acquire operative skills prior to completion of specialist training.[1] In addition, a culture change has lead to the expectation and demonstration of a certain level of competency prior to operative exposure. As a result, the emphasis of surgical training has shifted towards a proficiency-based system and the traditional teaching method of "see one, do one, teach one" has become outdated.[2]

TURP is a difficult procedure to master, with trainees expected to complete hundreds in order to achieve expertise. In recent years, developments in medical therapy have resulted in fewer TURP's being performed.[3,4] As a result, current trainees are unlikely to perform as many TURP's in their training when compared to their trainers.

Technological advances and improvements in computer graphics however, may offer an alternative to the traditional methods of surgical training. Virtual reality (VR) surgical simulators provide repetitive practice and performance feedback without requiring supervision in a safe environment.[5] Simulators have the potential to shorten the learning curve for complex surgical procedures, create skills which transfer to the operating room and therefore decrease the incidence of future complications.[6–8] In 2008 the American College of Surgeons proposed that all surgical training programmes should incorporate simulation.[9] However, before simulators can be adopted into the training curriculum they must be validated to ensure the instrument is a true representation of the real-life task in question (face validity) and that it can

* Corresponding author. Tel.: +44 01392 411611.
E-mail addresses: ebright@doctors.net.uk (E. Bright), S.Vine@exeter.ac.uk (S. Vine), Mark.Wilson@exeter.ac.uk (M.R. Wilson), mastersr@hku.hk (R.S.W. Masters), johnmcgrath@doctors.org.uk (J.S. McGrath).

1743-9191/$ – see front matter
doi:10.1016/j.ijsu.2012.02.012

accurately reflect differences in performance (construct validity). Furthermore, to be an effective educational tool, the simulator must improve key skills and provide targeted parameters for improvement.[10] The aim of this study is to assess the validity and training benefits of the TURPsim™ (Simbionix USA Corp., Cleveland, OH).

2. Method

18 participants volunteered to take part in the study (13 males, 5 females; mean age 35.8 years; range 20–60 years). Participants were both left ($n = 3$) and right ($n = 15$) hand dominant and consisted of 11 novices (non-urological junior doctors or medical students with no TURP experience) and 7 experts (>200 TURPs performed). All participants completed a median lobe prostate resection task using the TURPsim™.

The TURPsim™ uses a computer programme (VirtaMed AG, Zurich, Switzerland) to simulate a number of TURP-based resection tasks. These tasks range in difficulty from the resection of a single prostate lobe to a complete TURP, as well as other variables that increase the complexity of the resection task, such as bleeding. The simulator comprises a resectoscope and a docking station through which the resectoscope is passed, hence representing the urethra (see Fig. 1a and b). Once the resectoscope is fully inserted into the docking station, a computer generated intravesical view appears on the monitor. Withdrawal of the resectoscope brings the prostate and the urethra into view. The simulator is connected to a laptop computer on which specific training tasks are selected. Diathermy resection is achieved using foot pedals.

All participants completed "resection task 1", a simple task requiring resection of the median lobe only, during which no complications such as bleeding occur. This task was utilised in this study as it was felt appropriate for participants of all levels to attempt, as it did not require detailed knowledge of the anatomy of the prostate. All participants read a study-specific information sheet, provided written consent to participate and were given a brief demonstration of the basic controls and functions of the simulator. Participants watched a demonstration video of resection task 1 available on the TURPsim™, following which they were given the chance to familiarise themselves with the simulator controls by performing an unrecorded attempt at the task. All participants were instructed to stop resecting when they had reached a 70% resection target or 120 seconds had elapsed. Novice performers were then asked to complete a further four trials, with the same instruction and guidance.

For each of the recorded resection tasks, the following performance indicators were downloaded from the simulator; completion time, the percentage of prostate resected (PR), the percentage of capsular resection (CR) and the time that the diathermy loop was active without tissue contact (TAWC). To account for differences in completion time, a relative score was calculated (percentage resection ÷ completion time) indicating the amount of prostate resected per minute. This was used as a measure of general performance. Task safety was also calculated using a relative score (amount of time the diathermy loop was active without tissue contact ÷ completion time). The total percentage of capsular resection was taken directly from the simulator and used as a measure of performance error.

Verbal comments from the experts were obtained following completion of the resection task to assess face validity. Their opinion regarding the reality of the simulator and its appropriateness as a learning tool was sought.

Statistical analysis was performed using SPSS 19 software (SPSS Inc., Chicago, IL, USA). A series of independent sample *t*-tests were performed to assess differences between experts and novices for PR, CR and TAWC measures. A one-way repeated measures ANOVA was performed on all measures for the five learning trials performed by the novices. Significant effects were followed up with Bonferoni corrected post-hoc *t*-tests. For ANOVA tests effect sizes were calculated using partial eta squared (ηp2) for omnibus comparisons and for *t*-tests effect sizes were calculated using the Cohen's *d* method. A series of independent sample *t*-tests were then performed to assess differences between the novices' fifth attempt at the task and the experts' single attempt, for all measures.

3. Results

3.1. Experts vs un-trained novices

Experts resected a significantly greater percentage of prostate per minute (PR) ($M = 59.05$, SD $= 15.72$ vs $M = 30.86$, SD $= 4.78$; $t(16) = 5.64$, $p < 0.01$, $d = 2.43$) and had significantly less active diathermy time without tissue contact (TAWC) ($M = 0.19$, SD $= 0.33$ vs $M = 1.41$,SD $= 0.71$; $t(16) = 4.27$, $p < 0.01$, $d = 2.20$) than novices. Novices recorded a higher percentage of capsular resection (CR) than experts, although this difference did not reach statistical significance ($M = 13.71$, SD $= 5.71$ vs $M = 16.91$, SD $= 6.58$; $t(16) = 1.05$, $p = 0.31$, $d = 0.52$).

3.2. Novice training

ANOVA revealed that there were significant differences across the learning trials for all three measures; PR ($F(4,40) = 20.05$, $p < 0.01$. $\eta_p{}^2 = 0.667$), CR ($F(4,40) = 6.81$, $p < 0.01$, $\eta_p{}^2 = 0.405$) and TAWC ($F(4,40) = 12.55$, $p < 0.01$, $\eta_p{}^2 = 0.557$). Follow up *t*-tests revealed significant increases in PR between attempt 1 and 3 ($p < 0.05$), 1 and 4 ($p < 0.05$), 1 and 5 ($p < 0.01$), 2 and 4 ($p < 0.05$), 2 and 5 ($p < 0.01$) and 3 and 5 ($p < 0.01$) (see Fig. 2); significant reductions in CR between attempt 1 and 5 ($p < 0.05$) (see Fig. 2) and significant decreases in TAWC between attempt 1 and 5 ($p < 0.01$), 2 and 5 ($p < 0.01$) and 3 and 5 ($p < 0.01$) (see Fig. 3).

3.3. Experts vs. trained novices

Although novice PR increased with repeated training, experts still resected a greater amount of prostate compared to the novices fifth resection attempt ($M = 59.05$, SD 15.72 vs $M = 41.22$, SD $= 4.55$; $t(16) = 3.59$, $p < 0.01$, $d = 1.54$). However no significant

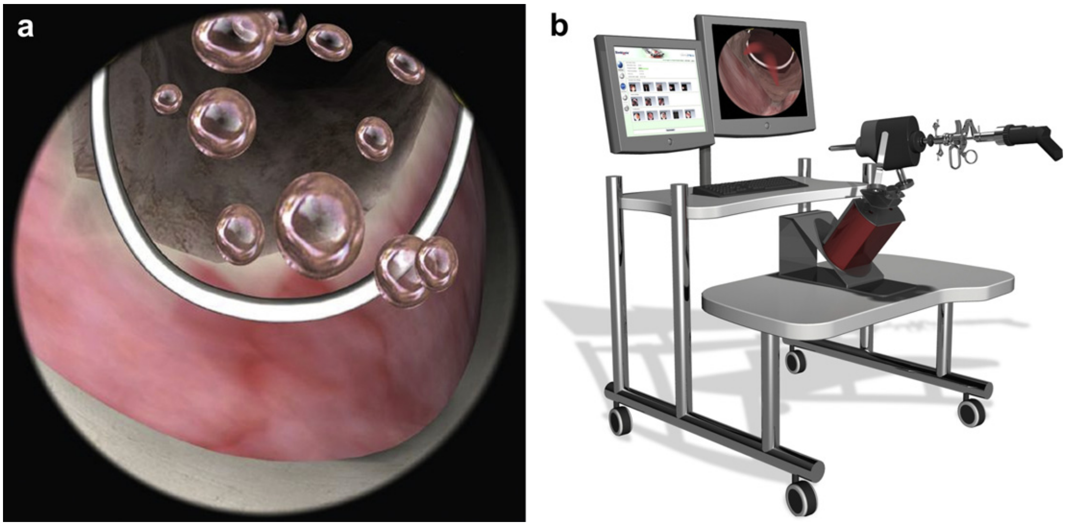

Fig. 1. (a) Example of a simulated median lobe prostate resection task. (b) Virtual reality TURPsim™ trainer (Simbionix USA Corp., Cleveland, OH).