Surg Endosc (2011) 25:166–171
DOI 10.1007/s00464-010-1151-6

European consensus on a competency-based virtual reality training program for basic endoscopic surgical psychomotor skills

Koen W. van Dongen · Gunnar Ahlberg · Luigi Bonavina · Fiona J. Carter · Teodor P. Grantcharov · Anders Hyltander · Marlies P. Schijven · Alessandro Stefani · David C. van der Zee · Ivo A. M. J. Broeders

Received: 21 January 2010 / Accepted: 18 May 2010 / Published online: 24 June 2010

Abstract
Background Virtual reality (VR) simulators have been demonstrated to improve basic psychomotor skills in endoscopic surgery. The exercise configuration settings used for validation in studies published so far are default settings or are based on the personal choice of the tutors. The purpose of this study was to establish consensus on exercise configurations and on a validated training program for a virtual reality simulator, based on the experience of international experts to set criterion levels to construct a proficiency-based training program.
Methods A consensus meeting was held with eight European teams, all extensively experienced in using the VR simulator. Construct validity of the training program was tested by 20 experts and 60 novices. The data were analyzed by using the *t* test for equality of means.
Results Consensus was achieved on training designs, exercise configuration, and examination. Almost all exercises (7/8) showed construct validity. In total, 50 of 94 parameters (53%) showed significant difference.
Conclusions A European, multicenter, validated, training program was constructed according to the general consensus of a large international team with extended experience in virtual reality simulation. Therefore, a proficiency-based training program can be offered to training centers that use this simulator for training in basic psychomotor skills in endoscopic surgery.

Keywords Endoscopic surgery · Virtual reality · Simulation · Training · Assessment · Education · Courses

K. W. van Dongen
University Medical Centre, Utrecht, The Netherlands

K. W. van Dongen
TweeSteden Hospital, Tilburg, The Netherlands

G. Ahlberg
Karolinska Hospital, Stockholm, Sweden

L. Bonavina
University of Milan, Milan, Italy

F. J. Carter
Yeovil District Hospital, Yeovil, UK

T. P. Grantcharov
Glostrup Hospital, Copenhagen, Denmark

T. P. Grantcharov
University of Toronto, St. Michael's Hospital, Toronto, Canada

A. Hyltander
Sahlgrenska Hospital, Goteborg, Sweden

M. P. Schijven
Academic Medical Centre, Amsterdam, The Netherlands

A. Stefani
Cisanello Hospital, Pisa, Italy

D. C. van der Zee
Wilhelmina Children's Hospital, Utrecht, The Netherlands

I. A. M. J. Broeders
Meander Medical Centre, Amersfoort, The Netherlands

I. A. M. J. Broeders (✉)
Twente University, Institute of Technical Medicine,
Drienerlolaan 1 Enschede (Noordhorst 120), PO box 217,
7500 AE Enschede, The Netherlands
e-mail: Iamj.broeders@meandermc.nl

Since the late nineties virtual reality (VR) simulators have been used to train residents in basic psychomotor skills for endoscopic surgery. Several studies have demonstrated a positive learning curve as well as improvement of skills in the operating room after training on these types of simulators [1–10]. Because these simulators allow accurate assessment of these skills, they may be used to define a skills level required for trainees to start surgical training in the actual operating room. At this moment only one study has been published on the development of an evidence-based VR training program for technical skills before progression into the operating room [11]. The optimal introduction of a VR simulator into an evidence-based, efficient, and cost-effective surgical skills curriculum is a core issue and is still open for discussion.

The LapSim® VR simulator basic skills module consists of nine psychomotor skill tasks. The end points relate to execution time, instrument path, damage, and other adverse effects. The training design can be adjusted for both exercise configurations and for "pass or fail" outcomes (assessment thresholds). The exercise configuration settings and exercise programs used for validation of this simulator in studies published so far are default settings or based on personal choice of the tutors [11–17]. Before claims about competency may be stated, consensus should be established on the settings of these exercise configurations and training programs. Second, such a training program should be constructed with expert performance as the guideline to define assessment thresholds.

The purpose of this study was to establish consensus on exercise configurations and on a training program for the LapSim® VR simulator, based on experience of international experts. Furthermore, the training program based on these configurations was validated to set criterion levels to construct a proficiency-based training program.

Methods

Equipment and tasks

The LapSim virtual reality simulator uses the Virtual Laparoscopic Interface (VLI) non-haptic-enhanced hardware platform (Immersion Inc., San Jose, CA), which includes a jig with two endoscopic handles. The VLI has an interface with a 2,600-MHz hyperthreading processor Pentium IV computer running Windows XP and is equipped with 256 RAM, GeForce graphics card, and 18-inch TFT monitor. The systems feature LapSim Basic Skills 3.0 software (Surgical Science Ltd, Göteburg, Sweden), from the LapSim Basic Skills package, comprises nine tasks for training basic psychomotor skills.

The computer stores and displays between 7 and 14 parameters of performance per task. These parameters are either *time-related* parameters (s), *error-related* parameters (e.g., tissue damage (mm and #), maximum stretch damage (0–100%), instrument misses (#), badly placed and dropped clips (#), blood loss (ml), rip failure (#), burn damage (#)) or *efficiency of instrument handling-related* (e.g., path length (m), angular path degree (°) and drift (mm)). Time, path length, and angular path degrees are measured for all nine tasks. The other parameters are measured subject to the nature of the task.

Training design and exercise configurations

To obtain an optimal structure for the training design and optimal configurations, a 2-day consensus meeting was organized hosting eight European teams, all extensively experienced in using the LapSim® VR simulator in surgical resident training programs. Participating centers to the consensus meeting were: Karolinska Hospital, Stockholm, Sweden; Glostrup Hospital, Copenhagen, Denmark; Surgical Skills Centre, Dundee, Scotland; Sahlgrenska University Hospital, Goteborg, Sweden; St. Mary's Hospital, London, United Kingdom; University of Milan, Milan, Italy; Cisanello Hospital, Pisa, Italy; University Medical Center, Utrecht, the Netherlands.

A structured questionnaire was used to determine the optimal training design (Table 1).

All experts proposed their personal exercise configurations. Consensus on preferable configuration was established by training and discussing all different configurations during a testing day.

Expert performance

A time frame of 6 months was set to test a maximum of five experts per center. The experts should at least have performed more than 100 endoscopic procedures and perform advanced endoscopic surgical procedures themselves

Table 1 Questionnaire to determine training and examination design

Which exercises should be used?
Modules with different difficulty?
If, yes, how many?
Which level should be examination level?
Thresholds based on expert scores?
If yes, mean score or mean + 1 × SD, mean + 2 × SD?
Maximum exposure time?
Should trainees pass exam once or more consecutive times?
Massed or distributive training?

SD standard deviation

(i.e., not perform basic endoscopic procedures defined as diagnostic laparoscopy, laparoscopic cholecystectomy, and laparoscopic appendectomy only).

Every task was started with a try out, which is considered to be a familiarization run to get comfortable with the simulator. Immediately after the familiarization run, a second run was performed to measure performance on a particular level. To avoid benefits resulting from immediate training, there had to be a break of 1 h between each level tested per expert. The results of the second run of experts were used to compare results of novices on the same level, to test for construct validity.

Novice performance

Every level of difficulty was tested with 20 novices. Because training on an easy or moderate level will lead to experience, novices were tested on one level only. Therefore, 60 novices were randomized, using the closed envelop method, into a training group for easy ($n = 20$), moderate ($n = 20$), or difficult ($n = 20$) levels. The novices—students or interns—had no previous experience performing or assisting in endoscopic surgery and displayed serious interest in a surgical career. When a parameter showed "construct validity" determined by a significant difference between the mean value of the experts and the mean value of the novice, this parameter was used as a threshold for the examination module.

Statistical analysis

The data were analyzed using SPSS (Chicago, IL) version 12.0.1., with the t test for equality of means. A power analysis on the data of van Dongen et al. [16] show that with a power of 0.8 and alpha set at 0.005, the sample size should be at least 17. Therefore, a minimal group size of 20 was chosen.

Results

Consensus on training design

The results of the 2-day consensus meeting are shown in this paragraph. In the "Discussion" section, the rationale for the consensus protocol is described.

The proposed training program exists of eight basic skill exercises on three different levels (easy, moderate, and difficult) and the suturing task. It was agreed on that trainees should start at the easy level, being tutored during their first familiarization run. If any of the parameters in this study show "construct validity," these parameters will be set as a threshold. Ultimately the "difficult level" of the training program also will be the accreditation level.

Consensus on exercise configurations

In general, targets will become smaller, disappear faster, or will be more vulnerable according to level of difficulty. The complete training schedule is available upon request.

Construct validity and thresholds

Six of the centers provided data of a total of 20 experts within the set timeframe of 6 months. The scores showed significant difference between novices and experts, and thus showing construct validity will be used as a threshold during training and examination of residents. The threshold is set at the mean score plus twice the standard deviation (SD) of the experts' scores. In total 50 of 94 parameters (53%) showed significant difference between the expert group and novice group scores in favor of the expert group. Table 2 shows the P values of these repeating parameters for which the experts perform better than novices. Time shows construct validity for all tasks. Efficiency of movement parameters validate in 18 of 32 cases (56%). A

Table 2 Significant parameters and *P* values of general parameters/task

Task	Significant parameters/total parameters	Time (L&R)	Path length (L&R)	Angular path (L&R)	Tissue damage	Maximum damage	Misses (L&R)
Camera navigation	2/7	**0.000**	0.3	0.3	0.2	0.2	**0.000**
Instrument navigation	7/10	**0.0/0.0**	0.13/**0.08**	0.5/0.4	**0.0**	**0.0**	**0.0/0.0**
Coordination	5/10	**0.0**	0.3/0.5	0.4/0.8	0.4	**0.01**	**0.0**
Grasping	6/10	**0.0/0.0**	0.5/0.4	1/0.3	**0.01**	**0.04**	**0.0/0.0**
Lift and grasp	7/9	**0.0**	**0.04**/0.2	**0.0/0.0**	**0.05**	0.06	**0.0/0.0**
Cutting	9/11	**0.0**	**0.0/0.0**	**0.0/0.0**	**0.0**	0.15	N.A.
Clip applying	7/10	**0.0**	0.6/**0.03**	0.8/**0.02**	N.A.	**0.04**	N.A.
Fine dissection	3/14	**0.01**	0.6/**0.08**	0.6/0.3	N.A.	N.A.	N.A.
Suturing	5/14	**0.01**	0.1/0.9	0.2/0.7	0.1	0.2	N.A.

(*L&R*) left and right; *N.A.* not applicable minority with 21 of 51 of the error scores validate (41%). Column 2 of Table 2 also shows the number of significant parameters in regard to the total number of parameters per task. Time, path length, and angular path were measured for every task. Tissue damage and maximum damage is measured for all tasks except for clip applying and fine dissection. The same accounts for misses, which also is not applicable to the cutting and suturing tasks.

Consensus on examination

All of the sessions (easy, moderate, and difficult) are to be trained up to threshold levels, based on the construct valid expert scores. When failing in three consecutive runs in any exercise, the trainee will be advised to continue to the next exercise and try this particular exercise again later. The trainees will have an official examination at the difficult level. The thresholds are based on the mean scores of the experts plus twice the SD. The trainees should pass these requirements twice. Of course the training sessions should be organized according to logistic possibilities of the training centers or hospitals, but distributed training is preferred. A maximum exposure time of 45 min per training session is advised. A training session will be stopped after 45 min or after finishing a level with success.

The trainees should pass the exercises within one training session; otherwise, they should perform the examination again at another time. Residents in training should ideally not be allowed to start with endoscopic surgery before passing this virtual reality training curriculum.

Discussion

A European, multicenter, validated, training program was constructed according to general consensus of a large international team with extended experience in virtual reality simulation.

Consensus on the training program

The LapSim® VR simulator is programmed with default training settings by the manufacturer; if desirable these settings can be adjusted according to buyer's preference. All training programs described in earlier studies are using default settings or are based on personal choice [11–19]. The default settings tend to be unrealistic at the difficult level [11]. The optimum training settings are not yet known and therefore we organized a meeting with eight European centers to reach consensus. These centers have all had extensive experience with control of settings for the LapSim® Virtual Reality Simulator. This experience served as the starting point to reach consensus on preferable exercise configuration and training design.

All eight basic skills were planned to be implemented into the training program, as well as the suturing task. It has been discussed if suturing should actually be a part of a basic skills training program. Although suturing can be considered a more demanding task, novices should be familiar with suturing before entering an operating room, to be able to solve minor complications if necessary.

Three different levels (easy, moderate, and hard) of difficulty were chosen to teach the residents the basic psychomotor skills gradually. To provide an efficient equipment familiarization, trainees should be properly instructed during the first training sessions by a tutor.

All centers reported that consecutively failing of a particular training task is likely to result in heightened levels of frustration among trainees. When failing a task three times in succession, frustration levels appear to influence the concentration level. Consequently, residents are to be advised to stop training a specific task after failing this task three times in a row.

A training session that exceeds 45 min in duration is likely to result in a decrease in concentration and accuracy and must be avoided. Furthermore, several studies have shown that distributed training, with rest between training sessions, is superior to massed training for obtaining psychomotor skills [20, 21]. Although within every hospital or training center training schedules will depend on local possibilities, distributed training with sessions of 45 min is advised. Moreover, a training session of the eight basic skills exercises can be finished within this timeframe.

Construct validity and thresholds

As shown in Table 2, the time-related and efficiency of instrument handling parameters very often show construct validity. Apart from the fine dissection task, time does always validate. The parameters that show efficiency of instrument handling validate very often with 21 of 32 parameters (65%) showing construct. These particular construct parameters are used to define the thresholds for a proficiency-based training program.

The suturing task shows construct validity in only 5 of 14 parameters. Most likely this is due to a lack of realism of the thread and the virtual tissue in this particular task. Consequently, the suturing task of this simulator cannot be considered as a validated training task assessing suturing skills. In contrast to these result, Munz et al. showed that there is transfer in skills from this suturing task to knot tying in a box trainer [22]. Twenty participants completed a correct knot compared with only five participants (25%) before training the suturing task. Time to completion was 66% faster and knot quality was 45% better after training.

Further research should be performed to show the benefit of the virtual reality suturing task in training endoscopic surgical skills.

The fine dissection task shows construct on a few parameters. At the easy and moderate levels, time is the only parameter that shows construct. Therefore, this exercise should not be a component of the training program, because it can be passed when performing very fast while causing a lot of damage.

Consensus on examination

The most challenging exercises are those with levels set at the most difficult level. Using this level, most parameters show construct validity. One may conclude that passing a test with parameter set at this level provides the best chance on appropriate basic skills in real surgery. Therefore, the difficult level is chosen as the accreditation level.

Schijven and Jakimowicz [23] have shown different performance profiles of trainees in acquiring basic skills. Some trainees need more training compared with others and therefore proficiency-based training is preferred. This means that residents should train until a preset level is reached, showing this basic level of proficiency. In our opinion the preset levels should be based on expert scores. Attaining basic psychomotor skills at an expert level could lead to shortened learning curves in patients. This, due to the fact that residents have attained technical basic psychomotor skills, combined with the fact that they might be able to perform surgery at a more automated level because they are less likely to “think” about their movements. During training in the operating room, residents will be able to focus on the procedure and decision making during surgery, instead of focussing on their psychomotor skills. Aggarwal et al. [24] bring up another advantage of an examination based on expert scores. They state that the confidence of residents in their own performance might be increased, having the knowledge that they are at least able to perform psychomotor skills at an expert level.

Trainees should show consistence in attained basic psychomotor skills by passing the examination twice. Passing the examination only once is more likely to be the outcome of chance.

The authors are aware that this simulator trains and measures trainees’ psychomotor skills solely. Because the acquisition of psychomotor skills has shown to decrease the learning curve in laparoscopic cholecystectomy [2, 10], this is, in our opinion, the minimum of objective assessment that should be performed before allowing residents to be active surgical team member in the operating room.

Even though endpoint parameters are defined scientifically and the training program settings are chosen based on consensus of international experts, this particular training program also should show predictive validity. Therefore, a study should be commenced to assess whether this training program predicts future performance and whether the trained skills transfer to the operating room.

Conclusions

A multicentered, European, broad, training program was constructed according to general consensus of eight LapSim® VR simulator user teams, who all had personal experience in validating this simulator. The results of this study define the parameters that can be utilized for the benchmark criteria of a training program. Therefore, a proficiency-based training program can be offered to training centers that use this simulator for training basic psychomotor skills in endoscopic surgery.

Acknowledgment The authors thank R. Aggarwal of the Imperial College of London for sharing his knowledge and for his collaboration during the consensus meeting.

Disclosures Drs. Koen van Dongen, Dr. Gunnar Ahlberg, Prof. Dr. Luigi Bonavina, Dr. Fiona Carter, Dr. Teodor Grantcharov, Dr. Marlies Schijven, Dr. Alessandro Stefani, Prof. Dr. David van der Zee, and Prof. Dr. Ivo Broeders have no conflicts of interest or financial ties to disclose. Dr. Anders Hyltander was founder and director of Surgical Science, Sweden, from 1999 until 2004. At this moment he is senior advisor to this company.

References

1. Aggarwal R, Ward J, Balasundaram I, Sains P, Athanasiou T, Darzi A (2007) Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. Ann Surg 5:771–779
2. Ahlberg G, Heikkinen T, Iselius L, Leijonmarck CE, Rutqvist J, Arvidsson D (2002) Does training in a virtual reality simulator improve surgical performance? Surg Endosc 1:126–129
3. Ahlberg G, Enochsson L, Gallagher AG, Hedman L, Hogman C, McClusky DA III, Ramel S, Smith CD, Arvidsson D (2007) Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. Am J Surg 6:797–804
4. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P (2004) Randomized clinical trial of virtual reality simulation for laparoscopic skills training. Br J Surg 2:146–150
5. Gurusamy K, Aggarwal R, Palanivelu L, Davidson BR (2008) Systematic review of randomized controlled trials on the effectiveness of virtual reality training for laparoscopic surgery. Br J Surg 9:1088–1097
6. Hyltander A, Liljegren E, Rhodin PH, Lonroth H (2002) The transfer of basic skills learned in a laparoscopic simulator to the operating room. Surg Endosc 9:1324–1328
7. Kundhal PS, Grantcharov TP (2009) Psychomotor performance measured in a virtual environment correlates with technical skills in the operating room. Surg Endosc 3:645–649
8. Larsen CR, Soerensen JL, Grantcharov TP, Dalsgaard T, Schouenborg L, Ottosen C, Schroeder TV, Ottesen BS (2009) Effect

of virtual reality training on laparoscopic surgery: randomised controlled trial. BMJ 14:b1802
9. Schijven MP, Jakimowicz JJ, Broeders IA, Tseng LN (2005) The Eindhoven laparoscopic cholecystectomy training course-improving operating room performance using virtual reality training: results from the first E.A.E.S. accredited virtual reality trainings curriculum. Surg Endosc 9:1220–1226
10. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, Satava RM (2002) Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg 4:458–463
11. Aggarwal R, Grantcharov TP, Eriksen JR, Blirup D, Kristiansen VB, Funch-Jensen P, Darzi A (2006) An evidence-based virtual reality training program for novice laparoscopic surgeons. Ann Surg 2:310–314
12. Duffy AJ, Hogle NJ, McCarthy H, Lew JI, Egan A, Christos P, Fowler DL (2005) Construct validity for the LAPSIM laparoscopic surgical simulator. Surg Endosc 3:401–405
13. Eriksen JR, Grantcharov T (2005) Objective assessment of laparoscopic skills using a virtual reality stimulator. Surg Endosc 9:1216–1219
14. Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J (2003) Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. Am J Surg 2:146–149
15. Sherman V, Feldman LS, Stanbridge D, Kazmi R, Fried GM (2005) Assessing the learning curve for the acquisition of laparoscopic skills on a virtual reality simulator. Surg Endosc 5:678–682
16. van Dongen KW, Tournoij E, van der Zee, Schijven MP, Broeders IA (2007) Construct validity of the LapSim: can the LapSim virtual reality simulator distinguish between novices and experts? Surg Endosc 21(8):1413–1417
17. Panait L, Bell RL, Roberts KE, Duffy AJ (2008) Designing and validating a customized virtual reality-based laparoscopic skills curriculum. J Surg Educ 6:413–417
18. Aggarwal R, Crochet P, Dias A, Misra A, Ziprin P, Darzi A (2009) Development of a virtual reality training curriculum for laparoscopic cholecystectomy. Br J Surg 9:1086–1093
19. Verdaasdonk EG, Dankelman J, Lange JF, Stassen LP (2008) Incorporation of proficiency criteria for basic laparoscopic skills training: how does it work? Surg Endosc 12:2609–2615
20. Verdaasdonk EG, Stassen LP, van Wijk RP, Dankelman J (2007) The influence of different training schedules on the learning of psychomotor skills for endoscopic surgery. Surg Endosc 2:214–219
21. Mackay S, Morgan P, Datta V, Chang A, Darzi A (2002) Practice distribution in procedural skills training: a randomized controlled trial. Surg Endosc 6:957–961
22. Munz Y, Almoudaris AM, Moorthy K, Dosis A, Liddle AD, Darzi AW (2007) Curriculum-based solo virtual reality training for laparoscopic intracorporeal knot tying: objective assessment of the transfer of skill from virtual reality to reality. Am J Surg 6:774–783
23. Schijven MP, Jakimowicz J (2004) The learning curve on the Xitact LS 500 laparoscopy simulator: profiles of performance. Surg Endosc 1:121–127
24. Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A (2006) A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. Am J Surg 1:128–133