Basic Laparoscopic Training Using the Simbionix LAP Mentor: Setting the Standards in the Novice Group

Martin W. von Websky, MD,* Martina Vitz, MD,† Dimitri A. Raptis, MD,* R. Rosenthal, MD,‡ P. A. Clavien, MD,* and Dieter Hahnloser, MD*

*Department of Visceral Surgery, University Hospital of Zurich, Zurich, Switzerland; †Laparoscopic Training Center, Zurich, Switzerland; and ‡Department of Surgery, University Hospital of Basel, Basel, Switzerland

BACKGROUND: Virtual reality devices are becoming the backbone for laparoscopic training in surgery. However, without knowledge of the achievable metrics of basic training within the trainee group, these simulators cannot be used effectively. Currently, no validated task metrics of the performance of larger trainee groups are available.

STUDY DESIGN: From April 2004 to December 2009, we collated an extensive prospective database using the Simbionix LAP Mentor (Simbionix USA, Cleveland, Ohio) for basic laparoscopic training of novice surgeons. This database was used to determine benchmarks for basic skill exercises and procedural tasks that combine stimulus to improve and feasibility with acceptance of the training program and the goal to train for safe surgery.

RESULTS: In all, 18,996 task performances of 286 novice trainees were analyzed. For the basic skill exercises, the total time for correct execution ranged between 45 seconds for basic skill 3 (eye-hand coordination) and 269 seconds for basic skill 9 (object placement). For the procedural tasks, the total time for correct execution ranged between 68 seconds for procedural task 1 (clipping and cutting) and 256 seconds for procedural task 3 (dissection). The total time to task completion depended mainly on right instrument path length with high correlation to left instrument path length. Learning curve analyses of the 4 procedural tasks demonstrated performance plateaus after 10–15 repetitions. Most complications occurred during the initial repetitions of the respective task. The best quartile of performances was chosen as peer group benchmark because it provides sufficient stimulus for improvement without discouraging trainees, thus enhancing adherence to the training program. The benchmark for safety and accuracy parameters was set at a predefined level of 95% correct execution.

CONCLUSIONS: As experience with virtual reality (VR) training is growing, curricula must be based on benchmarks for efficient training derived from large trainee groups to optimize use of the still costly simulators. Safety parameters should be included in trainee assessment. We share a set of metrics that take into account both performance and feasibility for basic laparoscopic training of surgical novices using the Simbionix LAP Mentor. (J Surg 69: 459-467. © 2012 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: laparoscopy, laparoscopic training, proficiency benchmarks, surgical education, patient safety, virtual reality

COMPETENCIES: Patient Care, Practice Based Learning and Improvement, Systems Based Practice

INTRODUCTION

Simulator training has been used increasingly to enhance and promote laparoscopic skills in surgical residency since the introduction of laparoscopic procedures in the 1980s.[1] The great advantage of virtual reality (VR) training devices is an objective and safe measurement of performance parameters while providing high face validity.[2] Evidence from the literature suggests that VR training is beneficial at least for novices to become familiar with the principles of laparoscopic surgery[3] without jeopardizing patient safety.[4] Evidence from randomized controlled trials indicates that VR training improves real operating room performance.[5-7] Even a short, preoperative "warm up" on a VR trainer has been shown to improve surgeons' performances.[8]

However, to use VR trainers effectively in surgical education, the benchmark metrics of the performed tasks must be determined first and then monitored along with the trainees' progress. To date, available data in the literature regarding most of the commonly used VR simulators are limited. The Simbionix LAP Mentor (Sim-

Correspondence: Inquiries to Dieter Hahnloser, MD, Department of Visceral Surgery, University Hospital of Zurich, Raemistrasse 100, CH 8091 Zurich, Switzerland; fax: +41 (0)44,255 44 49; e-mail: dieter.hahnloser@chuv.ch

Supported by Grant 32003B-120722 from the Swiss National Foundation.

bionix USA, Cleveland, Ohio)[9] is one of the most commonly used virtual training devices.[10-12] However, because no benchmark metrics for the different skill exercises have been defined as yet, no formative trainee assessment has been possible.

Whereas advanced laparoscopic training must aim for trainees to reach expert laparoscopic surgeon performance, basic laparoscopic training must be formative and feasible, and it must encourage novice surgeons to reach a sufficient basic standard in their peer group. The aims of this study were to determine benchmark criteria for the Simbionix LAP Mentor based on an extensive prospective data collection of surgical trainees without previous laparoscopic experience and to establish a valid data set enabling novices to lay the foundation for further successful laparoscopic training.

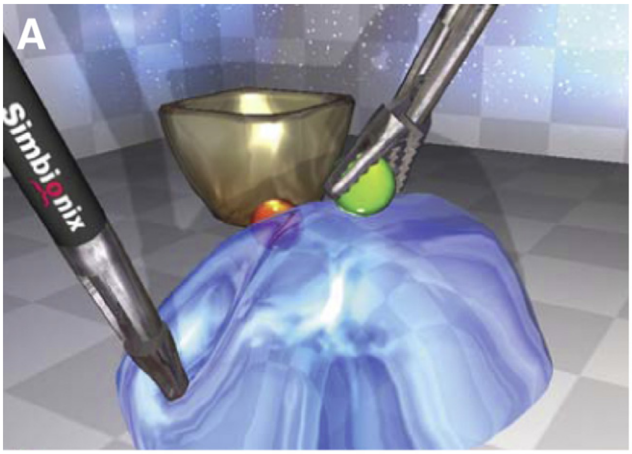
MATERIAL AND METHODS

Between April 2004 and December 2009, our group collected data of 286 surgical novices prospectively using the Simbionix LAP Mentor for laparoscopic training with trainees from Switzerland, Germany, Israel, Denmark, and Estonia. In total, 18,996 task performances of 9 basic skills and 5 procedural tasks were analyzed to set benchmarks indicating proficiency within the peer group. The study was funded by the Swiss National Foundation (SNF No: 32003B-120722). No financial support was received from Simbionix.

Users were first- to second-year residents, all novices in laparoscopic surgery, having performed less than 5 laparoscopic procedures, and without previous VR experience. The trainees were familiarized with the simulator according to a standardized protocol consisting of a 2-hour demonstration of all tasks and at least 1 supervised (1:1 supervision) performance of each task. Next, the trainees performed the exercises according to the assigned training course with prospective performance parameter recording. Subjects were told to repeat each basic skill and procedural task at least 3 times. Training for longer than 2 hours per session was discouraged. After each task a short break was mandatory.

Performances were analyzed for total time until task completion and several other parameters (eg, instrument path lengths, safety parameters, such as “safe cautery” and others). Next, a benchmark for each task consisting in a trade-off between quality of performance and feasibility had to be determined.

To be widely acceptable, this benchmark must be easy to calculate, yet also reflect the preceding mentioned compromise between stimulus to improve and avoidance of trainee discouragement. The authors attempted to solve this problem by taking into account the time and repetitions necessary for three quarters of the trainees to reach stable performance in the respective tasks. Thus, for each basic skill exercise and procedural task, we defined the threshold to the best 25% of performances of the measured parameters over all recorded sessions as “proficiency in the task” in our course. To emphasize our goal of training for safe surgery, we defined “proficiency” for safety and accuracy parameters arbitrarily (such as “safe cautery” as recorded by the Lap Mentor) when the respective score was above 95% correct execution. For procedural tasks 1-4, extensive learning curve analyses were performed to assess the necessary repetition number for stable performance. Examples of the VR environment are shown for 1 basic skill exercise and 1 procedural task, respectively, in Fig. 1.



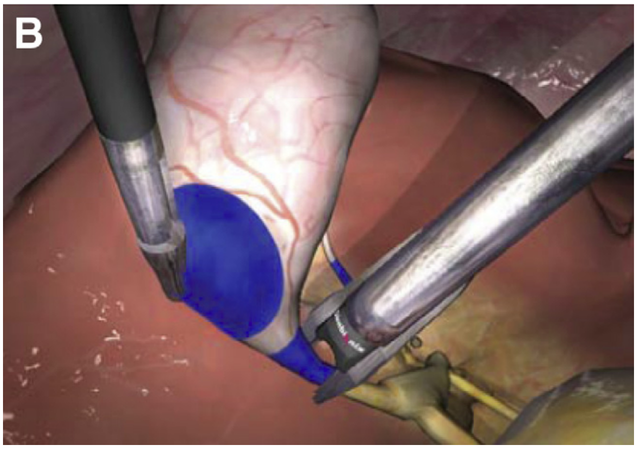


FIGURE 1. Examples of the VR environment for 1 basic skill and 1 procedural task. (**A**) Basic skill task 2. Two-handed maneuvers: improving bimanual skills, tissue handling, and hand–eye coordination by exposing the balls and placing them in the receptacle. (**B**) Procedural task 2. Clipping and cutting—2 hands. Exposing Calot’s triangle by correct gallbladder traction for safe cutting and clipping.

The data were collected in a prospective database and analyzed using all trainee data sets. The performance parameters were recorded by the Simbionix LAP Mentor software and output files were created using Microsoft Excel (Microsoft Corp, Redmond, Washington). A statistical analysis was performed using PASW statistics 18.0.2 (SPSS, IBM, Chicago, Illinois) using nonparametric tests. For basic skill and procedural task analysis, all available data were cross checked by 2 authors (M.W. and M.V.) for plausibility. The 25% percentile cut-offs (or 75% percentile cut-offs, where appropriate) representing the threshold to the best quartile of all performances of the defined variables were calculated.

For the 4 procedural tasks, a learning curve analysis was performed using Friedman’s test (nonparametric repeated-

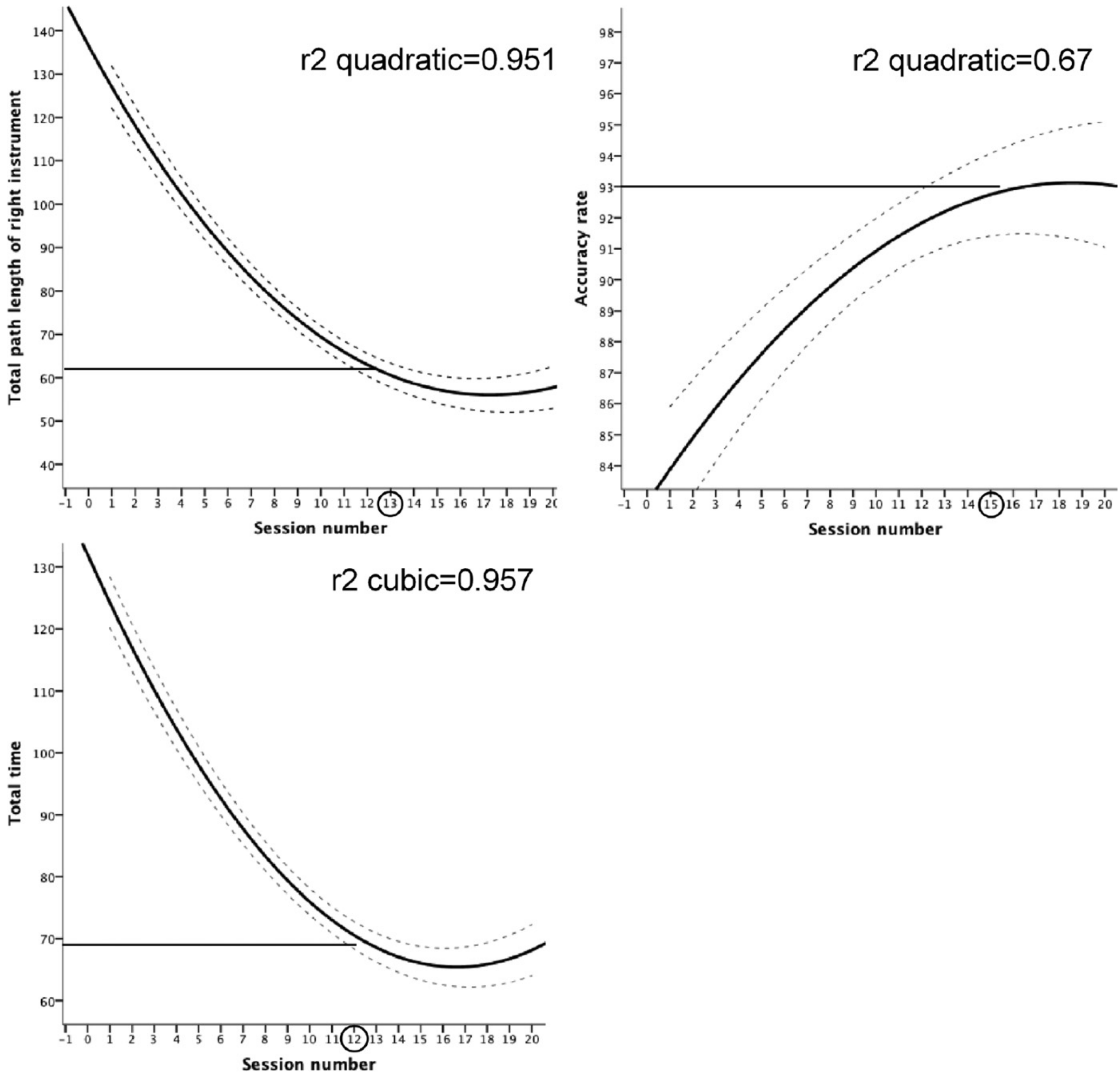


FIGURE 2. Learning curves of parameters in procedural task 1. The bold line represents the line best fitted to the data (line fitting method of least squares, see Reference 21), and the interrupted lines represent the standard error for each measurement. The function type is also given. For an easier comparison, the circle indicates at which repetition the plateau performance (horizontal line) is reached as calculated by the Dunnett's T3 correction (Table 3).

measures analysis of variance) to show that differences existed among the median performances in respect to the number of repetitions.[13-15] This analysis demonstrates only that substantial differences in performance according to session number exist; it does not demonstrate the trend of this difference (improvement or worsening of performance). Next, to determine the trend of the difference, multiple comparisons of individual task repetitions to identify when skills reached a plateau were performed using the Dunnett T3 correction.[16-20] A $p < 0.05$ was considered statistically significant. This analysis illustrates that the observed difference in performance is indeed a performance improvement. To visualize this finding and identify the respective learning curves, line fitting was performed for the measured parameters (Fig. 2). The underlying function types for each parameter (in this case mostly quadratic and cubic function types) were determined using http://eBiostatistics.com.[21] Curve fitting was based on the method of least squares using different curves (linear, exponential, logarithmic, and polynomial third and fourth degrees) for fitting.[22] In all cases, the coefficient of determination R^2, which measures the quality

TABLE 1. Performance Parameters and Best 25th Percentile Cut-Off in Basic Skill Exercises 1–9 on the Simbionix LAP Mentor

Basic Skill	Short Description of Skill	Parameter	Best 25th/75th Percentile, [Median] (Range)
1	Camera manipulation 0°: Repetitions n = 1416	Total time (s)	74 [85] (27–315)
		Maintaining horizon (%)	97.7 [94] (0–100)
		Total path length (cm)	141 [178] (34–955)
		Average speed of camera (cm/s)	<8.7 [9] (6–26)
2	Camera manipulation 30° Repetitions n = 1686	Total time (s)	82 [99] (67–340)
		Accuracy rate (%)*	100 [90] (0–100)
		Total path length (cm)	208 [269] (12–1377)
		Average speed of camera (cm/s)	<7.3 [9] (6–379)
3	Eye–hand coordination: Repetitions n = 2285	Total time (s)	45 [56] (34–149)
		Accuracy rate (%)*	100 [100] (20–100)
		Economy of movements (right) (%)	70 [60] (7–100)
		Economy of movements (left) (%)	74 [63] (9–883)
		Average speed (left/right) (cm/s)	(<1.9/<1.9) [2.4/2.4] (1/1.2–9.8/6.2)
4	Clip applying Repetitions n = 2284	Total time (s)	55 [64] (41–223)
		Economy of movement (right) (%)	64 [53] (7–100)
		Economy of movement (left) (%)	70 [56] (6–308)
		Average speed (left/right) (cm/s)	(<2.1/<2.2) [2.5/2.6] (1/1–9.5/14.5)
5	Clipping and grasping: Repetitions n = 1941	Total time (s)	96 [119] (16-315)
		Economy of movement (right) (%)	60 [49] (7-98)
		Economy of movement (left) (%)	62 [51] (7-86)
		Average speed (left/right) (cm/s)	(<2.8/<2.3) [3.3/2.7] (1.3/1.2-13.1/15.4)
6	2-handed maneuvers: Repetitions n = 1784	Total time (s)	97 [123] (5–610)
		Economy of movements (right) (%)	53 [42] (5–2796)
		Economy of movements (left) (%)	53 [42] (7–1684)
		Average speed (left/right) (cm/s)	(<2.4/2.3) [2.8/2.8] (1.2/1.2–7.9/6.3)
7	Cutting: Repetitions n = 1403	Total time (s)	103 [140] (28–964)
		Accuracy rate (%)*	100 [100] (0–100)
		Safe retraction (%)*	95 [50] (0–100)
		Average speed (left/right) (cm/s)	(<1.8/2.1) [2.1/2.5] (1/1.1-8.8/15)

TABLE 1. Continued.

Basic Skill	Short Description of Skill	Parameter	Best 25th/75th Percentile, [Median] (Range)
8	Coagulation: Repetitions n = 1197	Total time (s)	201 [281] (32–1280)
		Efficiency of cautery (%)*	95 [92] (43–100)
		Cut highlighted bands	21 [21] (1–21)
		Average speed (left/right) (cm/s)	(<1.7/1.6) [2/1.8] (1/1–12/4.7)
9	Object placement: Repetitions n = 809	Total time (s)	269 [356] (81–3783)
		Efficiency of translocation (%)*	100 [82] (9–100)
		Total path length (left/right) (cm)	810/979 [1088/1390] (81/242–11, 169/24,378)
		Average speed (left/right) (cm/s)	(<2.0/2.2) [2.3/2.5] (1.4/1.5–6.8/7.3)

*For safety parameters, >95% score was required.

of the fit, was calculated ($0 < R^2 < 1$). Next, the type of curve that best combined plausibility and goodness of fit (ie, coefficient of determination R^2 approaching 1) was chosen. Furthermore, Spearman's correlation coefficient was used to identify which of the defined parameters showed high correlation, with the aim of avoiding repetitive measuring of redundant parameters in future.

RESULTS

Of the trainees, 47% were men and 53% women, and 8% were left handed. The median trainee age was 27 years, ranging from 26 to 35 years. All were novices in laparoscopic surgery with less than 5 laparoscopic procedures and no previous VR experience. Video gaming experience further was present in a minority of the trainees (9.8%) and was not assessed in this study.

Basic skill tasks

For the basic skill exercise analysis, 14,805 task performances of 286 trainees were included. The 25th percentile cut-off (or 75th percentile cut-off where appropriate), representing the threshold to the best quartile of performances, defined proficiency within the peer group in the respective basic skill exercise. The parameter and the respective values are depicted in Table 1. As stated in the Material and Methods section, proficiency for safety and accuracy parameters (such as the "safe cautery" parameter as recorded by the Lap Mentor) were reached when the respective score was above 95% correct execution.

Procedural tasks

For the procedural task analysis 4,191 task performances of 245 trainees were included. The 25th percentile cut-off (or 75th percentile cut-off where appropriate), likewise defining proficiency within the peer group for each procedural task and parameter, are listed in Table 2. The test results for Friedman's test indicate that differences existed among the median performances in respect to the number of repetitions in all 4 procedural tasks (Friedman test: $p < 0.001$ in procedural tasks 1-4), thus demonstrating learning curve behavior.

Spearman's correlation analysis was carried out to unmask redundant parameters with high interparameter correlation. Spearman's nonparametric correlation showed mostly high correlation between the total time needed until task completion and right instrument path lengths (Spearman's correlation coefficient [CC] = 0.7 in task 1, CC = 0.8 task 2, CC = 0.9 tasks 3 and 4, respectively), indicating that total time for task completion depends mainly on the path length of the right instrument. Also, instrument path lengths left and right correlated well if 2 instruments were used (Spearman's CC = 0.7 in task 2; CC = 0.8 and CC = 0.9 in tasks 3 and 4, respectively).

Finally, curve fitting was performed to visualize the learning curves for each parameter in all 4 procedural tasks. As an example, fitted curves are shown for procedural task 1 in Figure 2. For all 4 procedural tasks, the median parameter values at the performance plateau and the number of repetitions until the performance plateaus of the measured parameters were reached—as determined by Dunnett's T3 correction—are reported in Table 3. For the total time, safety parameters, and path length of the instruments, the plateau was reached between the 10th and 13th repetition (also determined by Dunnett's T3 correction and curve fitting). One exception was "safe cutting" in procedural task 2, where the plateau was reached after the seventh repetition. The average instrument tip speeds did not show a learning curve-associated pattern. The learning curves derived from the described fitting method put out mostly quadratic and cubic functions, the latter consisting of an initial plateau and subsequent improvement only after much more

TABLE 2. Performance Parameters and Best 25th Percentile Cut-Off in Procedural Task Exercises 1–4 on the Simbionix LAP Mentor

Procedural Task	Short Description of Skill	Parameters	Best 25th/75th Percentile, [Median] (Range)
1	Clipping and cutting–retracted gallbladder Repetitions n = 1029	Total time (s)	68 [93] (10–611)
		Accuracy rate (%)*	100 [87] (13–100)
		Path length (cm)	61 [88] (3–875)
		Average speed (right) (cm/s)	<1.8 [2.1] (1–17)
2	Clipping and cutting–2 hands Repetitions n = 1390	Total time (s)	81 [115] (10–463)
		Safe cutting (mm)	>2.9 [2.1] (1–6.3)
		Path length (right) (cm)	52 [74] (1–366)
		Path length (left) (cm)	26 [40] (12–409)
		Average speed (left/right) (cm/s)	(<1.5/1.6) [1.8/1.9] (1/1–11/204)
3	Dissection–achieving the "Critical View" Repetitions n = 934	Total time (s)	256 [382] (14–2505)
		Safe cautery (%)*	95 [89] (17–100)
		Path length (right) (cm)	328 [484] (395–2260)
		Path length (left) (cm)	76 [130] (20–1262)
		Average speed (left/right) (cm/s)	(<1.4/1.6) [1.7/1.8] (0.7/1–36/100)
4	Gallbladder separation Repetitions n = 838	Total time s	164 [293] (13–2428)
		Efficiency of cautery (%)	90 [76] (9–100)
		Path length (right) (cm)	156 [291] (2–2304)
		Path length (left) (cm)	89 [181] (10–2502)
		Average speed (left/right) (cm/s)	(<1.7/1.5) [2/1.7] (0.9/0.9–9.3/21)

*For safety parameters >95% score was required.

training effort. Most complications as recorded by the Lap Mentor system occurred during the first 5 repetitions. For complications in procedural task 1, see Fig. 3.

Formative basic training course

Appendix A, which is available at http://www.lapcenter.ch, shows the proficiency criteria to be achieved by the trainees in each task combined into a formative basic training course.

Estimated time needed to achieve proficiency within the peer group

The theoretical minimal time to complete all basic skill and procedural task exercises was calculated. The calculation was based on performing all 9 basic skills 5 times (taking the quartile threshold of the fastest performances as the time needed for 1 repetition). The procedural tasks were accounted for as often as needed to reach the performance plateau at least thrice on 2 consecutive occasions (again, with the fastest quartile threshold as time needed for 1 repetition). Three-minute breaks between repetitions, 5-minute breaks between tasks, and 15-minute breaks after each hour of simulator training were also accounted for. With these considerations—using our benchmarks as goals—completion time for the basic skills and procedural tasks amounts to approximately 11 hours, without taking into account instruction time. Our experience shows that the proficiency levels can be reached by most trainees after 6-7 sessions of approximately 2 hours of training.

DISCUSSION

In this study, we present benchmarks for basic laparoscopic training on the Simbionix LAP Mentor using a large data pool

TABLE 3. Number of Repetitions Necessary Until Performance Plateau of Parameters is Reached as Determined by Dunnett's T3 Correction in Procedural Tasks 1–4 on the Simbionix LAP Mentor

Procedural Task	Value When Plateau Reached	Session Reaching Plateau*
1† Sessions n = 825		
Total time (s)	63	13
Accuracy rate (%)	93	15
Path length right (cm)	68	12
Average speed right (cm/s)	ns	ns
2† Sessions n = 973		
Total time (s)	90	11
Safe cutting (mm)	2.3	7
Path length left (cm)	38	11
Path length right (cm)	70	11
Average speed left (cm/s)	ns	ns
Average speed right (cm/s)	ns	ns
3† Sessions n = 701		
Total time (s)	290	11
Safe cautery (%)	ns	ns
Path length left (cm)	90	11
Path length right (cm)	375	11
Average speed left (cm/s)	ns	ns
Average speed right cm/s)	ns	ns
4† Sessions n = 670		
Total time (s)	220	10
Efficient cautery (%)	79	10
Path length left (cm)	130	10
Path length right (cm)	230	10
Average speed left (cm/s)	ns	ns
Average speed right (cm/s)	ns	ns

ns, not significant at the 0.05 level.

*Performed by the post hoc Dunnett's T3 correction; the session number indicates reaching a plateau.

†Friedman test <0.0001.

of 286 novices. Using these data for formative and summative trainee (self-) assessment will facilitate a comparison of trainee progress and performance. The necessary trade-off between feasibility and quality of performance was achieved by using the best 25% of performances as benchmark cut-off for most tasks and parameters. While other benchmark values might prove equally useful, the upper quartile combined with our emphasis on safe surgery practices (95% correct execution) allows for easy application and results in high adherence to the training program, at the same time avoiding trainee frustration. Learning curve analyses for the 4 procedural tasks indicate that the necessary repetitions for maximum training effect were feasible within a reasonable time frame.

The benefits of proficiency level based VR training using Surgical Science software (Surgical Science Inc, Minneapolis, Minnesota) have been proven,[23] and they had a significant impact on real operating room performance.[24] However, the analysis provided by Ahlberg et al.[24] was based on only 13 trainees in each group. Furthermore, Crochet et al. have been able to demonstrate that VR training improves dexterity in inexperienced surgeons and leads to skill transfer onto a porcine model.[25] These are but 2 of many examples of skill transfer to real life surgery after VR practice. In the aviation industry and in military training, virtual reality trainers have been used for decades with significant and measurable success. Construct validity for the Simbionix Lap Mentor has been demonstrated,[10] and in Europe, a consensus exists on the use of competency-based VR training programs for basic laparoscopic training.[26] However, for most simulators, useful benchmarks for basic laparoscopic training are not available. Apart from the costs involved, this could be a reason why VR training has not found the widespread acceptance it deserves. Therefore, the main intention of this article is to provide surgical trainers with the means of comparing their individual trainees with the data of our large peer group when using the LAP Mentor (or a similar device) and to assess the trainee's progress up to the point from which it is sensible to move on to more sophisticated exercises and/or real life surgery. We, among many stakeholders, are convinced that basic training on the LAP Mentor most likely also results in a real-life improvement in the operating room. However, to prove this beyond doubt, an efficient basic training with formative assessment must be established first for each VR system. This is what we propose in this study after evaluating the performances of 286 novices using the Simbionix Lap Mentor. Basic-level training metrics must be generated from a large pool of users so that interindividual differences in psychomotor skills have only a negligible effect. We believe it is necessary first to derive the benchmarks for basic training from a peer group of trainees rather than from experts to provide a feasible yet challenging basic training before more advanced training modules, where expert performances have to be met, are attempted. In most testing environments, including medical examinations, students are always tested first within the peer group and not against expert performance. Arguably, often this creates unrealistic and hardly feasible training curricula that may result in trainee discouragement. The formative training course based on our findings is intended as a feasible first basic course before progressing to more advanced laparoscopic simulation, where expert performance will have to be met. Using our proposed benchmarks, this formative training course can be completed in approximately 11 hours, split into 6-7 2-hour sessions. Thus, the aim of creating the desired training effect without trainee frustration can be achieved. Others have agreed on the potential benefit of such data sets and benchmarks for efficient training and simulator use.[27]

VR simulators measure many different parameters that can be confusing to both trainer and trainee. In this study, we tried to limit the benchmark parameters to the most reliable ones, taking into account not only the relevant literature[10,24,28-30] and analytic considerations (learning curve behavior) but also safety parameters. In our opinion, safety and accuracy parameters should be given high priority (eg, over fast execution) even in basic laparoscopic training. Hence, the safety and accuracy parameters in our course were defined to be executed correctly (at a 95% rate) for proficiency in the respective task to avoid training effects aiming only at completion time while neglecting safe techniques. In our group of novice surgeons, serious

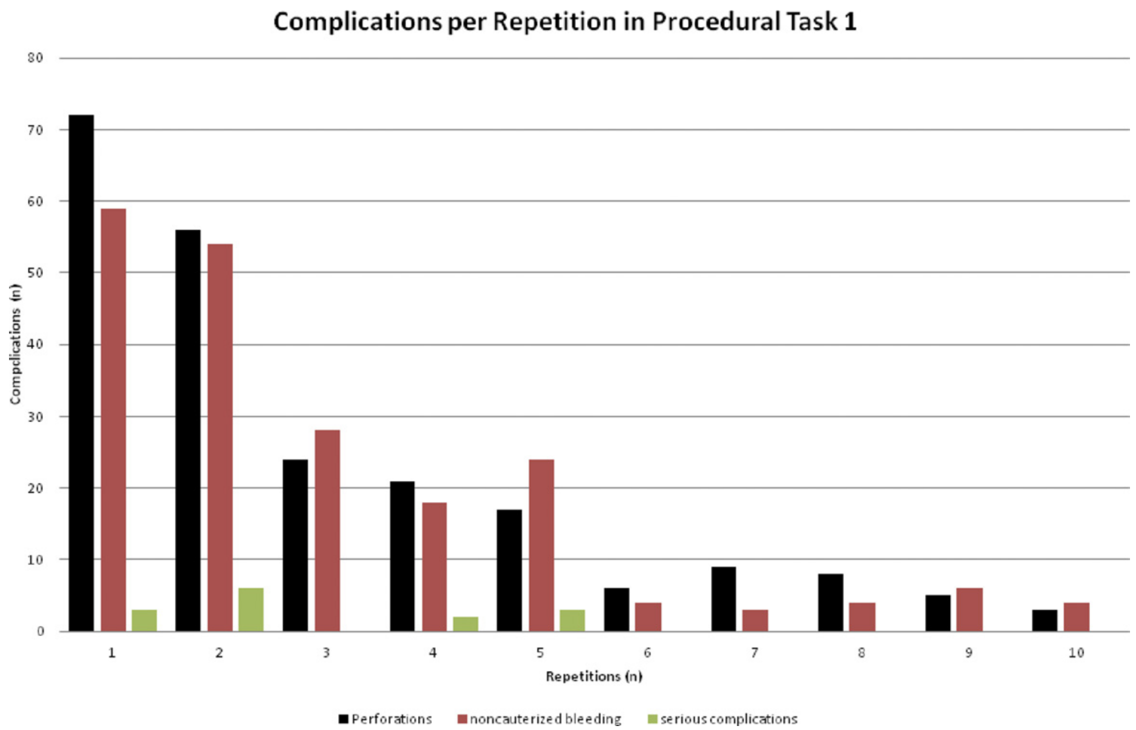


FIGURE 3. Complications in procedural task 1 as recorded by the Lap Mentor system depicted against the number of repetitions. The number of perforations, noncauterized bleeding, and serious complications are illustrated, respectively.

complications were most abundant in the initial performances of each task. Therefore, simulator training based on peer group benchmarks as proposed here, even when consisting only of 6-7 brief training sessions of 2 hours each, may serve as a valuable tool to avoid initial severe complications of novice laparoscopic surgery. It should, therefore, be recommended for each novice before attempting more sophisticated laparoscopic training.

As with many manual tasks, learning curve patterns were identified for most parameters. These learning curves differed in their form with cubic and quadratic functions most abundant. However, 1exemption from the learning curve pattern was instrument tip speed. This is an expected finding because neither excessive nor very low speed is a criterion for good performance. We, therefore, do not recommend using the instrument tip speed parameter as a quality assessment parameter for trainee performance.

Not every trainee will reach a level of satisfactory performance after a fixed number of repetitions for each task. Thus, it might not be advisable to prescribe a fixed number of repetitions when designing training curricula. In this study, maximum benefit was achieved after 10-15 training sessions for most novices, although some might need more or fail to reach the training goals.[31] To avoid trainee frustration, supervision, and assessment at different intervals by an instructor is imperative.[32]

In summary, as experience with VR training is growing, basic training curricula must be based on benchmarks to make best use of the still costly simulators. Safety parameters must be included in trainee assessment. After reaching the basic benchmarks provided in this study (see Appendix A, available also with additional data and learning curves at http://www.lapcenter.ch), the trainees should be able to progress quickly to advanced laparoscopic training where expert benchmarks must be met.

ACKNOWLEDGMENTS

The authors thank Jan Kukleta, MD, Christian Oberkofler MD, and Jêrome Gapany for their support. We also thank the Frieda-Locher-Hofmann Foundation for their engagement to establish laparoscopic simulator training for young surgeons.

REFERENCES

1. Mühe E. Long-term follow-up after laparoscopic cholecystectomy. *Endoscopy*. 1992;24:754-758.
2. 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.
3. Ganai S, Donroe JA, St Louis MR, Lewis GM, Seymour NE. Virtual-reality training improves angled telescope skills in novice laparoscopists. *Am J Surg*. 2007;193:260-265.
4. Kohn L, Corrigan J, Donaldson M, eds. *To Err Is Human: Building a Safer Health System*. Washington, DC: Institute of Medicine, National Academies Press; 1999.

5. Grantcharov TP, Kristiansen VB, Bendix J, et al. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg*. 2004;91:146-150.

6. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg*. 2002;236:458-463; discussion: 463-454.

7. Andreatta PB, Woodrum DT, Birkmeyer JD, et al. Laparoscopic skills are improved with LapMentor training: results of a randomized, double-blinded study. *Ann Surg*. 2006;243:854-860; discussion: 860-853.

8. Calatayud D, Arora S, Aggarwal R, et al. Warm-up in a virtual reality environment improves performance in the operating room. *Ann Surg*. 2010;251:1181-1185.

9. Symbionix Inc. Available at: http://www.simbionix.com/index.html.

10. Aggarwal R, Crochet P, Dias A, et al. Development of a virtual reality training curriculum for laparoscopic cholecystectomy. *Br J Surg*. 2009;96:1086-1093.

11. Lucas SM, Zeltser IS, Bensalah K, et al. Training on a virtual reality laparoscopic simulator improves performance of an unfamiliar live laparoscopic procedure. *J Urol*. 2008;180:2588-2591; discussion: 2591.

12. Zhang A, Hünerbein M, Dai Y, Schlag PM, Beller S. Construct validity testing of a laparoscopic surgery simulator (lap mentor): evaluation of surgical skill with a virtual laparoscopic training simulator. *Surg Endosc*. 2008; 22:1440-1444.

13. Siegel S, Castellan NJ. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, Inc; 1988.

14. Conover WJ. *Practical Nonparametric Statistics*. 2nd edition. New York: John Wiley and Sons; 1980.

15. Daniel WW. *Biostatistics*. 6th edition. New York: John Wiley and Sons; 1995.

16. Brown MB, Forsythe AB. Robust tests for the equality of variances. *J Am Stat Assoc*. 1974;69:364-367.

17. Levene H. Robust tests for the equality of variance. In: Olkin I, ed. *Contributions to Probability and Statistics*. Palo Alto, Calif: Stanford University Press; 1960.

18. Milliken G, Johnson D. *Analysis of Messy Data*. vol 1. *Designed Experiments*. New York: Chapman & Hall; 1992.

19. Neter J, Wassermann W, Kutner MH. *Applied Linear Statistical Models*. 3rd edition. Boca Raton, Fla: CRC Press; 1990.

20. Welch BL. On the comparison of several mean values: an alternative approach. *Biometrika*. 1951;38:330-336.

21. Function approximation. Available at: http://eBiostatistics.com. Accessed October 10, 2010.

22. Wolberg J. Data analysis Using the Method of Least Squares: Extracting the Most Information from Experiments. 1st edition. Berlin: Springer; 2005

23. Surgical Science. Available at: http://www.surgical-science.com/.

24. Ahlberg G, Enochsson L, Gallagher AG, et al. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *Am J Surg*. 2007;193:797-804.

25. Crochet P, Aggarwal R, Dub SS, et al. Deliberate practice on a virtual reality laparoscopic simulator enhances the quality of surgical technical skills. *Ann Surg*. 2011;253: 1216-22.

26. van Dongen KW, Ahlberg G, Bonavina L, et al. European consensus on a competency-based virtual reality training program for basic endoscopic surgical psychomotor skills. *Surg Endosc*. 2011;25:166-171.

27. Sugden C, Aggarwal R. Assessment and feedback in the skills laboratory and operating room. *Surg Clin North Am*. 2010;90:519-533.

28. Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. *Am J Surg*. 2006;191:128-133.

29. Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. *Am J Surg*. 2003;185:146-149.

30. Patel AD, Gallagher AG, Nicholson WJ, Cates CU. Learning curves and reliability measures for virtual reality simulation in the performance assessment of carotid angiography. *J Am Coll Cardiol*. 2006;47:1796-1802.

31. Grantcharov TP, Funch-Jensen P. Can everyone achieve proficiency with the laparoscopic technique? Learning curve patterns in technical skills acquisition. *Am J Surg*. 2009;197:447-449.

32. Stefanidis D, Acker CE, Swiderski D, Heniford BT, Greene FL. Challenges during the implementation of a laparoscopic skills curriculum in a busy general surgery residency program. *J Surg Educ*. 2008;65:4-7.