

Development of a virtual reality training curriculum for phacoemulsification surgery

AV Spiteri¹, R Aggarwal², TL Kersey³, M Sira⁴,
 L Benjamin^{5,6}, AW Darzi⁷ and PA Bloom⁸

¹London Kent Surrey Sussex Deanery, London, UK

²Division of Surgery, Department of Surgery and Cancer, Imperial College London, London, UK

³Hillingdon Hospital, Western Eye Hospital, London, UK

⁴Queen Alexandra Hospital, Portsmouth, UK

⁵Stoke Mandeville Hospital, Buckinghamshire, UK

⁶Education Committee of the Royal College of Ophthalmologists, London, UK

⁷Imperial College, Imperial College Hospital NHS Trust, Royal Marsden Hospital, Institute of Cancer Research, London, UK

⁸The Hillingdon Hospital NHS Foundation Trust and The Western Eye Hospital (Imperial College NHS Trust), Imperial College School of Medicine, Middlesex University, London, UK

Correspondence:
 AV Spiteri, Ophthalmology Department, Kings College Hospital, 92 Beaconsfield Road, EN3 6AP, Middlesex, London, UK.
 Tel: +44 (0)7818232255;
 Fax: +44 (0)20 3299 3445.
 E-mail: avspiteri@gmail.com

Received: 20 December 2012
 Accepted in revised form:
 19 August 2013
 Published online: 27 September 2013

Abstract

Purpose Training within a proficiency-based virtual reality (VR) curriculum may reduce errors during real surgical procedures. This study used a scientific methodology to develop a VR training curriculum for phacoemulsification surgery (PS).

Patients and methods Ten novice-(n) (performed <10 cataract operations), 10 intermediate-(i) (50–200), and 10 experienced-(e) (>500) surgeons were recruited. Construct validity was defined as the ability to differentiate between the three levels of experience, based on the simulator-derived metrics for two abstract modules (four tasks) and three procedural modules (five tasks) on a high-fidelity VR simulator. Proficiency measures were based on the performance of experienced surgeons.

Results Abstract modules demonstrated a 'ceiling effect' with construct validity established between groups (n) and (i) but not between groups (i) and (e)—Forceps 1 (46, 87, and 95; $P < 0.001$). Increasing difficulty of task showed significantly reduced performance in (n) but minimal difference for (i) and (e)—Anti-tremor 4 (0, 51, and 59; $P < 0.001$), Forceps 4 (11, 73, and 94; $P < 0.001$). Procedural modules were found to be construct valid between groups (n) and (i) and between groups (i) and (e)—Lens-cracking (0, 22, and 51; $P < 0.05$) and Phaco-quadrants (16, 53, and 87; $P < 0.05$). This was also the case with Capsulorhexis (0, 19, and 63; $P < 0.05$) with the performance decreasing in the (n) and (i) group but improving in the (e) group (0, 55, and 73; $P < 0.05$) and (0, 48, and 76; $P < 0.05$) as task difficulty increased.

Conclusion Experienced/intermediate benchmark skill levels are defined allowing the development of a proficiency-based VR

training curriculum for PS for novices using a structured scientific methodology.

Eye (2014) 28, 78–84; doi:10.1038/eye.2013.211; published online 27 September 2013

Keywords: phacoemulsification surgery; virtual reality; training curriculum

Introduction

Training in the operating theatre is often unstructured, and occurs by chance encounters dependent on patient and disease variability. A particular facet of surgical practice is the need to train inexperienced individuals to a level of competence in their chosen field. Although training is supervised, and in accordance with the informed consent of the patient, this probably may no longer be an ethically or economically viable option for modern medical practice. It is thus necessary to explore, define, and implement modes of surgical skills training that do not expose the patient to preventable errors.¹

There are many tools currently available for training and assessment in phacoemulsification surgery (PS) outside the theatre.² Laboratory practice allows surgeons to acquire skills in a controlled environment, free of the pressures of operating on real patients according to Piaget's and Vygotsky's pedagogical philosophy of 'learning by doing'. Wet labs use cadaveric human or animal models, or synthetic eyes (designed specifically for performing phacoemulsification) to rehearse the steps of cataract extraction. However, these methods have been criticised for being unrealistic³ with inaccurate simulation of tissue consistency and anatomy⁴ and also lacking any form of objective assessment. Simulation in the form of virtual

reality (VR) and synthetic models have been proposed for technical skills training at the early part of the learning curve in other fields of surgery.^{5–8} VR simulators are now starting to be introduced as an adjunct to microsurgical skills courses. It is however considered as preferable for training to be structured within a standardized curriculum.⁹ This should constitute knowledge-based learning, a stepwise technical skills pathway, on-going feedback and progression towards proficiency goals, enabling transfer to the real environment.¹⁰

The aim of this study was to develop an evidence-based and stepwise VR training curriculum for acquisition of technical skills for PS. Although simulators have been evaluated as a part-task training platform for differentiating and developing basic ophthalmic microsurgical skills,^{11,12} this is the first time that a phacoemulsification simulator has been subjected to a structured scientific method for curriculum development.

Materials and methods

The study recruited subjects, divided into novice (performed fewer than 10 PSs), intermediate (50–200 PSs), and experienced (>500 PSs) operators. Recruitment was solely through personal communication. The only exclusion criterion was previous training experience with a phacoemulsification simulator.

The EYESI surgical simulator (VR magic, Mannheim, Germany) phacoemulsification (PS) interface was used for this study and both abstract skills (such as forceps training; Figure 1a) and procedural tasks (eg, capsulorhexis; Figure 1b) were assessed. The full PS procedures vary in terms of difficulty and the nine selected tasks reflect this. A detailed description of the selected simulator tasks is provided in Table 1.

Each of the four abstract skill and five procedural tasks were performed for two sessions by all novice, intermediate, and experienced subjects. All sessions were completed at least 1 h apart. Before commencing each task, every subject was provided with a full demonstration by an experienced operator and a one-on-one simulator familiarization session during which no assistance was provided.

Data for each of the performed tasks were measured objectively by the VR simulator inbuilt scoring software and comprised 14–31 metrics depending on the task. The data were transferred to the Microsoft Excel spreadsheet (Microsoft Corporation, Redmond, WA, USA).

Performance evaluation

Construct validity is a test of whether a model can differentiate between different levels of experience, and thus be used to assess performance.¹⁰ Comparison of median performance among the three groups of surgeons was used to assess whether each simulated task was construct valid and substantiated the use of the defined settings of the simulator to assess phacoemulsification technical skill.

The definition of *benchmark criteria* to be achieved before progression to the next stage of the curriculum was by calculation of the median score for each parameter during the second session for all experienced surgeons.

The way novices advanced through these clearly defined steps through comparative measurement of simulator derived metrics, that is construct validation, and benchmark definition, enabled the assembly of a *curriculum* for abstract and procedural training, based on the data rather than supposition. This provided an evidence- and proficiency-based pathway for novice surgeons to follow.

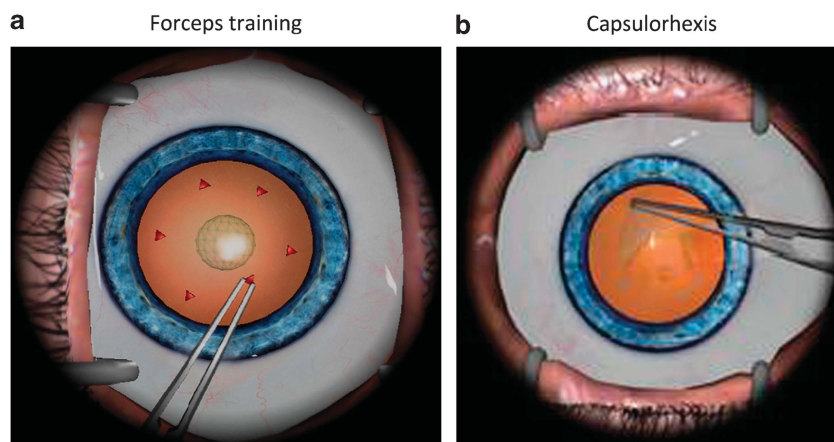


Figure 1 Screen shots of (a) forceps training abstract task and (b) capsulorhexis procedural task.

Table 1 Description of the selected modules on the EYESI VR phacoemulsification simulator

<i>Skill or task</i>	<i>Description</i>
<i>Abstract skills</i>	
Module 1—Cataract Forceps Training Module	
Task 1. Cubes with no gravity	Grasping 6 cubes with forceps and moving them from peripheral anterior chamber into central suspended sphere. Colour change indicating completion. No gravity
Task 2. Triangles with gravity	Grasping 6 triangles with forceps and moving them from peripheral anterior chamber into central suspended sphere. Colour change indicating completion. Gravitational tendency for triangles to drop towards lens
Module 2—Cataract Anti-Tremor Training	
Task 3. Straight path	Transferring ball on contact with rod along a straight line guide across anterior chamber
Task 4. Circular path	Transferring ball on contact with rod along a circular guide line around peripheral anterior chamber
<i>Procedural skills</i>	
Module 3—Capsulorhexis	
Task 5. Capsulorhexis using circle guide Tendency to run—low	Viscoelastic insertion. Initiation of rhexis with cystotome. Change instrument to forceps to complete continuous curvilinear capsulorhexis. Low tendency to radial extension
Task 6. Capsulorhexis using circle guide Tendency to run—medium	As task 5. Medium tendency to radial extension
Task 7. Capsulorhexis using circle guide Tendency to run—high	As task 5. High tendency to radial extension
Module 4—Phaco divide and conquer module	
Task 8. Cracking of four grooved quadrants	Insertion of phaco probe and second instrument to crack and rotate four grooved quadrants
Task 9. Phaco of four cracked quadrants	Setting of phaco power and aspiration settings. Insertion of phaco probe and second instrument phacoemulsify four cracked quadrants

Statistical analysis

The choice of 10 subjects per group was based on a two-tailed test, with $\alpha = 0.05$ and power $(1-\beta) = 0.80$, and an intended reduction of 30% in time taken to complete tasks for experienced *vs* novice operators, based on the data from previous studies of VR simulation.^{13,14} This yielded a value of eight subjects per group, which was increased to ten to allow for dropout and technical malfunction of the simulator.

The data were analysed with SPSS version 18.0 (SPSS, Chicago, IL, USA) using non-parametric tests. Comparison of performance between experienced, intermediate, and inexperienced groups was undertaken using the Kruskal–Wallis test (where $P < 0.050$ was considered as statistically significant) and the Mann–Whitney *U* test ($P < 0.017$ using the Bonferroni adjustment), as appropriate.

Results

Thirty subjects, comprising ten novices (n), ten intermediate (i), and ten experienced (e) operators, were recruited. All subjects completed two sessions on the four abstract skills, and two sessions on the five

procedural tasks. Only the statistically significant metrics selected will be discussed here. There was no statistically significant difference between the first and second repetition of these metrics except where otherwise stated. The second session scores were used for analysis to further reduce the effect of participant familiarization with the simulator during the first session. Construct validity was initially established overall for the total global scores of the nine selected tasks (Figure 2).

Abstract tasks

The metrics within the easier abstract modules demonstrated a ‘ceiling effect’ with construct validity established between (n) and (i) and between (n) and experienced (e) groups, but not between (i) and (e) groups.

Statistical significance was achieved primarily on global score—Anti-tremor 1 revealed a significant difference only in the first repetition and is excluded. Forceps 1 was significantly different between (n) and (i) (46, 87, and 95; $P < 0.001$ between (n) and (i)). Increasing difficulty of task showed a significantly reduced performance in global score in (n) but

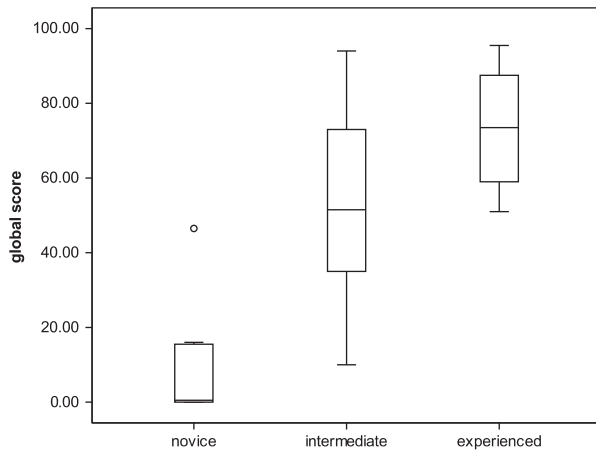


Figure 2 Total no. of global scores. Horizontal lines within boxes, boxes, and whiskers represent median, interquartile range, and range, respectively. Circle represents an outlier ($P \geq 0.05$, Kruskal–Wallis test).

minimal difference between (i) and (e)—Anti-tremor 4 (0, 51, and 59) and Forceps 4 (11, 73, and 94) both $P < 0.001$ between (n) and (i). Anti-tremor 1 and 4 showed similar results for average tremor value (47.1, 34.4, 34.3, and 45.6, 35.9, and 35.3; $P < 0.017$ between (n) and (i)).

Incision stress value in both tasks at both levels of difficulty also exhibited statistically significant differences between (n) and (i) but not between (i) and (e)—Anti-tremor 1 and 4 (0.23, 0, and 0; $P < 0.017$) and (3.4, 0, and 0; $P < 0.017$) and Forceps 1 and 4 (3.05, 0.03, and 0; $P < 0.017$) and (7.43, 0.12, and 0; $P < 0.017$).

Likewise, time taken in seconds exhibited significant differences between (n) and (i) only for the more difficult tasks Anti-tremor 4 and Forceps 4 (76.5, 54, and 52.5; $P \leq 0.017$) and (115.5, 71, and 68; $P \leq 0.017$) but not the easier Anti-tremor 1 and Forceps 1. This metric again demonstrated a ‘ceiling effect’ with the experienced group.

Procedural tasks

Procedural modules were found to be construct valid between groups (n) and (i) and between groups (i) and (e). This was the case for global score metrics in Lens cracking (0, 22, and 51; $P < 0.017$) and Phaco of quadrants (16, 53, and 87; $P < 0.017$). In capsulorhexis 1, the global scores demonstrated a similar trend (0, 19, and 63; $P < 0.017$). As the difficulty of the task increased (capsulorhexis 3 and 5), the global score performance in the (n) and (i) group decreased but improved in the (e) group (0, 55, and 73; $P < 0.017$) and (0, 48, and 76; $P < 0.017$).

Table 2 Construct valid metrics for task-capsulorhexis, level of difficulty-5, repetition-2

Metric	Construct valid	Benchmark level (median of experienced 2nd repetition)
Global score	✓	76
Time (s)		
Time (pts)		
Time with instruments (s)		
Incision stress (value)		
Incision stress (pts)		
Average radius of capsule (value)		
Radius deviation from reference 2.5 mm (value)	✓	0.03
Radius deviation from reference 2.5 mm (points)	✓	13.5 ^a
Decentration (value)		1.30 ^a
Decentration (points)	✓	1.28 ^a
Overall irregularity (value)	✓	
Overall irregularity (points)	✓	
Local irregularity spikes (value)	✓	0 ^a
Local irregularity spikes (points)	✓	0 ^a
Maximum radial extension (value)	✓	0.31
Maximum radial extension (points)	✓	94.1 ^a
Lens damage (value)	✓	3.185
Injured cornea area (value)		
Injured cornea area (points)		
Non-horizontal instruments insertion (number of events)		
Non-horizontal instruments insertion (points)		
Open forceps instrument insertion (number of events)		
Open forceps instrument insertion (points)		

^a Metrics not included in the curriculum (Figure 3).

In addition, in the capsulorhexis module, the more difficult the task performed, the more the number of significant metrics observed. Capsulorhexis 1 revealed a significant difference only for global score whereas capsulorhexis 5 had the most construct valid metrics (Table 2). However these, like the abstract tasks, exhibited statistically significant differences only between (n) and (i) but not between (n) and (e). These included Radial deviation value (0.18, 0.06, and 0.03; $P < 0.017$), Maximum radial extension value (3.33, 1.52, and 0.31; $P < 0.017$), and Lens damage value (15.95, 3.175, and 3.185; $P < 0.017$).

Curriculum construction

The statistically significant metrics only were used in the development of the training curriculum. The summarized outcome is a proficiency-based VR curriculum for training in PS (Figure 3).

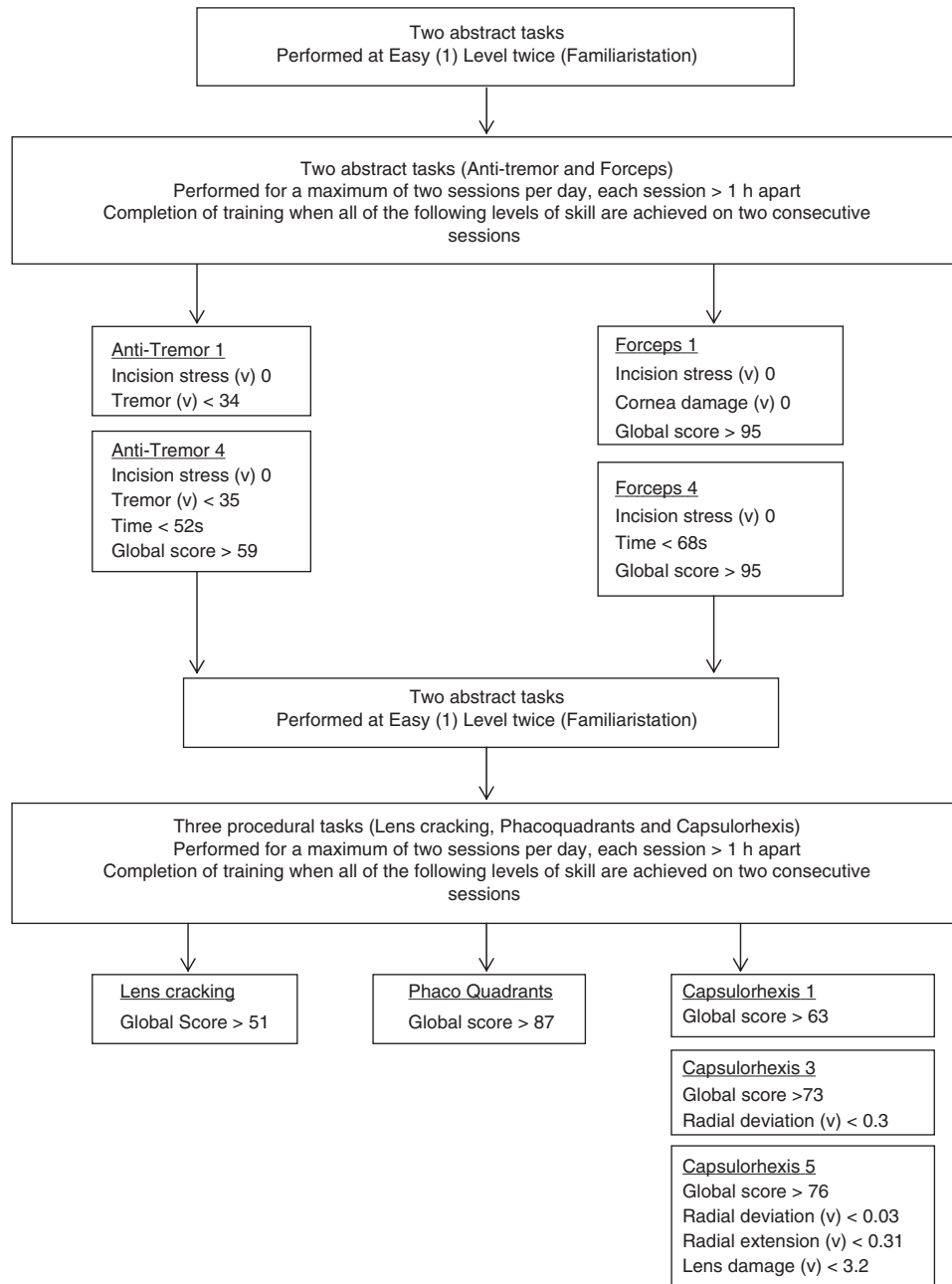


Figure 3 Evidence-based virtual reality training curriculum for PS (v) = value; (s) = seconds.

Discussion

This study applied a stepwise process to the modules and metrics of a VR simulator, resulting in the development of a training curriculum for PS. The modules were deemed to be construct valid through comparison of performance across three levels of surgical experience. Interestingly, there was no difference between the performance of intermediate and experienced groups on all the abstract tasks and on some of the procedural tasks on the simulator. This is an entirely appropriate finding, as those in the intermediate

group approach the plateau phase of their learning curve for PS. Inexperienced subjects are thus most likely to benefit from this training curriculum.

Training within the curriculum commences at the abstract skills modules, with two repetitions of all four skills. Progression to the procedural tasks necessitates achievement of the benchmark proficiency criteria, which are based on the scores derived from the performance of experienced phacoemulsification surgeons. The structure of the curriculum is identical for the five procedural tasks, which again have proficiency criteria for the

trainee to achieve before completion of the training period. It is also important to note that the curriculum adheres to the concept of 'distributed' rather than 'massed' training schedules, with a maximum of two sessions performed per day, each at least 1 h apart.^{15,16} Finally, to confirm acquisition of skill rather than attainment of a good score by chance, all benchmark levels must be achieved at two consecutive sessions.

Other studies have investigated the construct validity on the EYESi Phacoemulsification VR simulator. Mahr and Hodge¹⁷ as well as Le *et al*¹⁸ analysed performance of the anti-tremor and forceps modules while Privett *et al*¹⁹ focused on the easy and medium levels of the capsulorhexis module. However a fundamental for use of simulation in clinical training schedules, the organization of such data into a coherent, stepwise, and proficiency-based training curriculum has not yet been pursued.

A common complaint is the expense in terms of simulator cost and upkeep, training space and faculty time required for integration of VR curricula into residency programmes.²⁰ With reductions in the learning curve during real operations, it is possible that the total cost of training each surgical resident will be reduced. In terms of training schedules, this curriculum prescribes two sessions per day, at least 1 h apart. The evidence for distributed training schedules is clear, although it is uncertain whether this means practice once a day or once a week.^{11,12} Flexibility in accommodating training sessions will be needed when implementing this curriculum, but this should not detract from acquisition of skill as curriculum completion is based on the achievement of proficiency measures.

This training programme is not intended as a substitute for skills acquisition in the operating theatre, but it will allow part of the learning curve to be transferred to the skills laboratory.³ A cognitive skills module is also essential at the front end of any training programme, such as that available from the Royal College of Ophthalmologists microsurgical skills course. Furthermore, completion of this curriculum is based on the dexterity, rather than safety scores or clinical outcome measurements. It is important to use technical skills rating scales and to integrate such scales into the simulator software.²

It is crucial to disseminate this curriculum to other users of VR simulation, to enable external validation of the curriculum in terms of ease of use and feasibility, and to define learning curves to determine the minimum amount of repetitions necessary to establish proficiency. One could then ultimately confirm whether its use does actually lead to the notion of the pretrained novice who can operate with greater dexterity and skill on patients undergoing phacoemulsification surgical procedures.

It is then only a matter of time until other domains of ophthalmic surgical practice have to follow this lead of simulation-based training, with objective measurement of performance before operative intervention.

Summary

What was known before

- Evidence-based virtual reality (VR) training programmes have been used successfully in training novice surgeons in other fields of surgery. Some smaller ophthalmology studies have established construct validity for a limited number of tasks on the VR simulator, mostly comparing novices with experienced surgeons.

What this study adds

- To the best of our knowledge, this is the first evidence-based training curriculum for novice phacoemulsification surgeons using the VR simulator, with benchmark levels set by intermediate and experienced surgeons. Construct validity has been established more comprehensively for a number of abstract and procedural tasks selected to reflect most of the stages of phacoemulsification surgery. This was carried out for three different levels of experience defined by the surgical training courses run by the Royal College of Ophthalmologists (UK). Participant numbers (NNT) were determined by VR training validation trials successfully implemented in other fields of surgery.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

Emily Turton MSc (Research Assistant) for contributing towards statistical analysis. All running expenses funded privately.

This study has been previously presented at: The European Society of Cataract and Refractive Surgery (ESCRS) Annual Congress, Vienna, Austria on the 20th September 2011 where it was Winner of one of the Best of the Best at the Special Free Paper Session. The United Kingdom and Ireland Society of Cataract and Refractive Surgery (UKISCRS) Annual Congress, Southport, England, 14 October 2011* where it was voted Winner of the Best Cataract Surgery Paper. The study is being piloted at Regional Deanery level with a view to being implemented nationwide by the Royal College of Ophthalmologists as part of the training curriculum for all novice ophthalmic surgeons in the United Kingdom.

References

- 1 Aggarwal R, Darzi A. Technical-skills training in the 21st century. *N Engl J Med* 2006; **355**: 2695–2696.

- 2 Spiteri A, Aggarwal R, Kersey T, Benjamin L, Darzi A, Bloom P. Phacoemulsification skills training and assessment. *Br J Ophthalmol* 2010; **94**: 536–541.
- 3 Spitz L, Kiely EM, Piero A et al. Decline in surgical training. *Lancet* 2002; **359**: 83.
- 4 Smith JH. Teaching phacoemulsification in US ophthalmology residencies: can the quality be maintained? *Curr Opin Ophthalmol* 2005; **16**: 27–32.
- 5 Aggarwal R, Ward J, Balasundaram I, Sains P, Athanasiou T, Darzi A. Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. *Ann Surg* 2007; **246**: 771–779.
- 6 Aggarwal R, Tully A, Grantcharov T, Larsen CR, Miskry T, Farthing A et al. Virtual reality simulation training can improve technical skills during laparoscopic salpingectomy for ectopic pregnancy. *BJOG* 2006; **113**: 1382–1387.
- 7 Colg HG, Crawford SW, Galbraith 3rd O. Virtual reality bronchoscopy simulation: a revolution in procedural training. *Chest* 2001; **120**: 1333–1339.
- 8 Aggarwal R, Black SA, Hance JR, Darzi A, Cheshire NJ. Virtual reality simulation training can improve inexperienced surgeons' endovascular skills. *Eur J Vasc Endovasc Surg* 2005; **31**(6): 588–593.
- 9 Anastakis DJ, Wanzel KR, Brown MH, McIlroy JH, Hamstra SJ, Ali J et al. Evaluating the effectiveness of a 2-year curriculum in a surgical skills center. *Am J Surg* 2003; **185**: 378–385.
- 10 Aggarwal R, Grantcharov TP, Darzi A. Framework for systematic training and assessment of technical skills. *J Am Coll Surg* 2007; **204**: 697–705.
- 11 Solverson DJ, Mazzoli RA, Raymond WR, Nelson ML, Hansen EA, Torres MF et al. Virtual reality simulation in acquiring and differentiating basic ophthalmic microsurgical skills. *Simul Healthc* 2009; **4**(2): 98–103.
- 12 Feudner EM, Engel C, Neuhaan IM, Petermeier K, Bartz-Schmidt KU, Szurman P. Virtual reality training improves wet-lab performance of capsulorhexis: results of a randomized, controlled study. *Graefes Arch Clin Exp Ophthalmol* 2009; **247**(7): 955–963.
- 13 Aggarwal R, Grantcharov TP, Eriksen JR, Blirup D, Kristiansen VB, Funch-Jensen P et al. An evidence-based virtual reality training program for novice laparoscopic surgeons. *Ann Surg* 2006; **244**: 310–314.
- 14 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.
- 15 Mackay S, Morgan P, Datta V, Chang A, Darzi A. Practice. Distribution in procedural skills training: a randomized controlled trial. *Surg Endosc* 2002; **16**: 957–961.
- 16 Moulton CA, Dubrowski A, Macrae H, Graham B, Grober E, Reznick R. Teaching surgical skills: what kind of practice makes perfect? a randomized, controlled trial. *Ann Surg* 2006; **244**: 400–409.
- 17 Mahr MA, Hodge DO. Construct validity of anterior segment anti-tremor and forceps surgical simulator training modules: attending versus resident surgeon performance. *J Cataract Refract Surg* 2008; **34**(6): 980–985.
- 18 Le TD, Adatia FA, Lam WC. Virtual reality ophthalmic surgical simulation as a feasible training and assessment tool: results of a multicentre study. *Can J Ophthalmol* 2011; **46**(1): 56–60.
- 19 Privett B, Greenlee E, Rogers G, Oetting TA. Construct validity of a surgical simulator as a valid model for capsulorhexis training. *J Cataract Refract Surg* 2010; **36**(11): 1835–1838.
- 20 MacRae HM, Satterthwaite L, Reznick RK. Setting up a surgical skills center. *World J Surg* 2008; **32**: 189–195.