Ready for OR or not? Human reader supplements Eyesi scoring in cataract surgical skills assessment.

Selvander, Madeleine; Åsman, Peter

Published in:
Clinical Ophthalmology

DOI:
10.2147/OPTH.S48374

2013

Clinical Ophthalmology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

CASE SERIES

Ready for OR or not? Human reader supplements Eyesi scoring in cataract surgical skills assessment

Madeleine Selvander[1,2]
Peter Åsman[1]

[1]Department of Clinical Sciences, Malmö: Ophthalmology, Lund University, Malmö, Sweden; [2]Practicum Clinical Skills Centre, Skåne University Hospital, Malmö, Sweden

Purpose: To compare the internal computer-based scoring with human-based video scoring of cataract modules in the Eyesi virtual reality intraocular surgical simulator, a comparative case series was conducted at the Department of Clinical Sciences – Ophthalmology, Lund University, Skåne University Hospital, Malmö, Sweden.

Methods: Seven cataract surgeons and 17 medical students performed one video-recorded trial with each of the capsulorhexis, hydromaneuvers, and phacoemulsification divide-and-conquer modules. For each module, the simulator calculated an overall score for the performance ranging from 0 to 100. Two experienced masked cataract surgeons analyzed each video using the Objective Structured Assessment of Cataract Surgical Skill (OSACSS) for individual models and modified Objective Structured Assessment of Surgical Skills (OSATS) for all three modules together. The average of the two assessors' scores for each tool was used as the video-based performance score. The ability to discriminate surgeons from naïve individuals using the simulator score and the video score, respectively, was compared using receiver operating characteristic (ROC) curves.

Results: The ROC areas for simulator score did not differ from 0.5 (random) for hydromaneuvers and phacoemulsification modules, yielding unacceptably poor discrimination. OSACSS video scores all showed good ROC areas significantly different from 0.5. The OSACSS video score was also superior compared to the simulator score for the phacoemulsification procedure: ROC area 0.945 vs 0.664 for simulator score ($P = 0.010$). Corresponding values for capsulorhexis were 0.887 vs 0.761 ($P = 0.056$) and for hydromaneuvers 0.817 vs 0.571 ($P = 0.052$) for the video scores and simulator scores, respectively.The ROC area for the combined procedure was 0.938 for OSATS video score and 0.799 for simulator score (P=0.072).

Conclusion: Video-based scoring of the phacoemulsification procedure was superior to the innate simulator scoring system in distinguishing cataract surgical skills. Simulator scoring rendered unacceptably poor discrimination for both the hydromaneuvers and the phacoemulsification divide-and-conquer module. Our results indicate a potential for improvement in Eyesi internal computer-based scoring.

Keywords: simulator, training, cataract surgery, ROC, virtual reality

Correspondence: Madeleine Selvander
Department of Clinical Sciences, Skåne University Hospital, 101 Södra Förstadsgatan, Malmö S-205 02, Sweden
Tel +46 40 338 163
Fax +46 40 336 212
Email madeleine.selvander@med.lu.se

Background

Training with surgical simulation is becoming an important part of resident surgical training in ophthalmology.[1] A shift from counting cases to competence-based

curricula for learning cataract surgery is on its way, and the implementation of structured surgical curricula has also been shown to have a favorable impact on complication rates.[2,3] In many departments, simulation training is a required part of the training curriculum. Often, the simulator is also used for assessment, and residents are quantitatively evaluated on the surgical simulator as a form of diagnostic tool for cataract surgical skills.[1] Currently there are two commercially available virtual reality surgical simulators for cataract surgery: PhacoVision (Melerit Medical, Linköping, Sweden) and Eyesi (VRmagic, Mannheim, Germany). During the last 5 years, the Eyesi intraocular surgical simulator has been acquired by several institutions around the world. It was the predominant simulator used, according to a recent survey regarding the role of simulators in ophthalmic residency training in the US.[1] The Eyesi simulator has been evaluated for construct validity.[4,5] It has also been shown that training with the simulator improves wet-lab capsulorhexis performance.[6] One report also suggests an associated improvement in surgical performance on real cataract operations.[7] However, using the simulator for feasible assessment demands adequate scoring that allows the distinguishing of cataract surgical skill. This study therefore investigates the Eyesi as a diagnostic tool for cataract surgical skill by comparing the ability to distinguish cataract surgeons from nonsurgeons using either the internal Eyesi computer-based scoring or human video-based scoring of cataract modules in the Eyesi surgical simulator. To our knowledge, this is the first report evaluating a simulator with regard to performance score using receiver operating characteristic (ROC) curve analysis.

Materials and methods

Seventeen medical students and seven cataract surgeons were recruited to the study. None of them had prior experience with the Eyesi surgical simulator. The students were attending their seventh semester, and the simulator training took place during their ophthalmology rotation. The cataract surgeons worked at the university hospital or at a local hospital in the region. The group included five very skilled surgeons with over a thousand to several thousand cataract operations, and two junior surgeons with 20 and 150 self-reported cataract operations.

Simulator

The simulator investigated in this study was the Eyesi surgical simulator (VRmagic software version 2.5). The cataract head was used. The simulator consists of a model eye connected to a computer. Probes are inserted, and cameras inside the eye detect the movement of the probes. A virtual image is created on the computer. The virtual image is projected on two oculars, creating a binocular virtual image of the anterior segment. The image is also shown on an observer screen. Several manipulating and procedure-specific modules are available in the simulator. For this study, three cataract modules representing procedure-specific parts of a cataract operation were chosen: capsulorhexis, hydromaneuvers, and phacoemulsification (phaco) divide and conquer. In the capsulorhexis module (Figure 1), the trainee has to inject viscoelastics into the anterior chamber, create a flap of the anterior capsule with a cystotome, and create a capsulorhexis using a forceps. Here, level four out of ten was judged the appropriate level of difficulty most representing a real capsulorhexis procedure. In the hydromaneuvers module (Figure 1), level one out of four was judged to be the appropriate level. Here, the trainee has to create a visible fluid wave in the cortex and thereafter rotate the nucleus, showing that the hydrodissection is appropriate. For the phaco divide-and-conquer module (Figure 1), via the divide-and-conquer

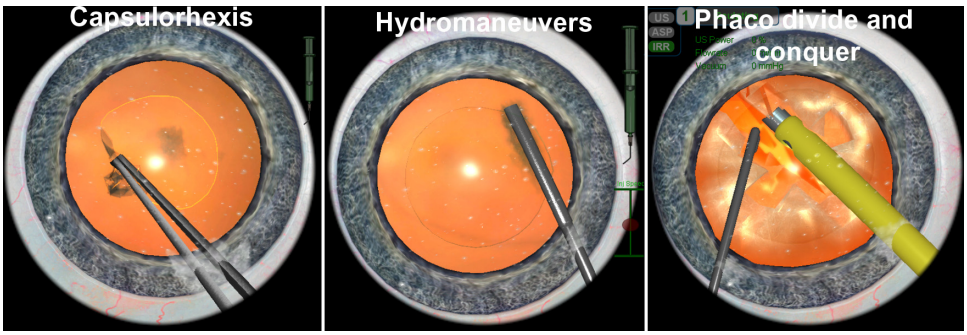

Figure 1 In the capsulorhexis procedure, the trainee has to inject viscoelastics into the anterior chamber, create a flap with a cystotome in the anterior capsule, and finally complete a circular capsulorhexis. In the hydromaneuvers module, the trainee places a cannula under the rhexis edge and injects liquid solution at the right speed, creating a visible fluid wave in the cortex. Afterwards, the trainee rotates the nucleus, proving that an appropriate hydrodissection has occurred. During the phacoemulsification (phaco) divide and conquer, the trainee has to create grooves in the nucleus, crack the nucleus into quadrants, and finally consume the quadrants with ultrasonic energy using phacoemulsification.

technique, the trainee has to create grooves in the nucleus, crack the nucleus into four quadrants, and finally remove each quadrant with phaco. Level five out of six was used.

A performance score ranging from 0 to 100 points is calculated by the simulator. The score is based on the efficiency of the procedure, target achievement, instrument handling, and tissue treatment.

Each participant performed three trials with each of the capsulorhexis and hydromaneuvers modules followed by two trials with the phaco divide-and-conquer module, in that order. For the purpose of this study, only the second trial was used for analysis and video recorded.

Video evaluation

The training session as seen on the observer screen was video-recorded. The films were randomized and evaluated by two experienced cataract surgeons who were blinded to the identity of the trainees. For evaluation of the saved video films, the Objective Structured Assessment of Cataract Surgical Skill (OSACSS) tool was used.[8] As previously described, this tool has been used to evaluate real cataract operations as well as video recordings from simulator training sessions.[9] The films were also evaluated with the modified Objective Structured Assessment of Surgical Skills (OSATS) tool, which has also been used for evaluation of real operations and simulator films.[10–12] The evaluations render each procedure a video-evaluation score where a higher score signifies higher surgical skills. Using OSACSS for capsulorhexis, the score lies between 3 and 15, for the hydromaneuvers, the score lies between 1 and 5, and for phaco divide and conquer, the score lies between 5 and 25. Using OSATS, all three procedures – capsulorhexis, hydromaneuvers, and phaco divide and conquer – are evaluated together, rendering scores that range between 4 and 20.

Statistical analysis

ROC curves were calculated for the simulator score and for the OSACSS and OSATS scores. The ROC curves for the OSACSS and OSATS scores were compared with the internal computer-based scoring, and the difference was evaluated using an algorithm suggested by DeLong et al.[13] Interrater-reliability scoring of the OSACSS and OSATS, respectively was analyzed using intraclass correlation coefficients.

Results

ROC-curve areas and the corresponding ROC curves are shown in Figure 2. The ROC area did not differ from 0.5 (random guess) for the simulator score for the modules hydromaneuvers and phaco divide and conquer. The ROC areas for video evaluations differed from 0.5 on all modules. See Figure 2 for *P*-values.

ROC curve-area values were higher for the video evaluations than for the simulator score, significantly so for the phaco divide-and-conquer module ($P = 0.01$). The differences for the other modules were not significant (capsulorhexis $P = 0.056$, hydromaneuvers $P = 0.052$).

Calculating the ROC curve for the combined procedure of capsulorhexis, hydromaneuvers, and phaco divide and conquer, where the scores from the individual procedures were added, yielded an ROC-curve area of 0.799 for the simulator scoring and 0.938 for the video-based scoring by OSATS. This difference was nonsignificant ($P = 0.072$).

Interrater correlation coefficients were high for the OSACSS capsulorhexis ($r = 0.788$), OSACSS phaco divide and conquer ($r = 0.726$), and OSATS ($r = 0.764$), and moderate for the OSACSS hydromaneuvers module ($r = 0.598$).

Discussion

The introduction of eye-surgery simulators for cataract-surgery training has been beneficial. Previous studies have shown construct validity for the Eyesi surgical simulator in the anterior segment.[4,5,12] Our results further support these findings.

It is advantageous that training can take place in a safe and standardized environment. Surgical simulators provide these conditions. Evaluation and assessment is often used in simulation-based training, where the trainee has to meet preset target criteria.[14–16] We decided to test the simulator as a diagnostic tool for cataract surgical skill by investigating the ROC curves. We found that the simulator scoring was weak on the hydromaneuvers and phaco divide-and-conquer modules, and comparing the ROC area for simulator score with 0.5, ie, random guess or flipping coins, rendered unacceptably poor discrimination for both modules. In contrast, the ROC areas using video-based scoring showed high values, indicating good diagnostic performance. One problem, especially with the phaco divide-and-conquer module scoring, was that the window of success was small. Therefore, even a small mistake could render a zero score. On the other hand, instruments could also be handled inappropriately with little effect on the score.

A limitation in our material is that our group of cataract surgeons was quite small, and represents a large variation in cataract surgical skills. Removing the two junior surgeons improved the ROC curves for both simulator scoring and

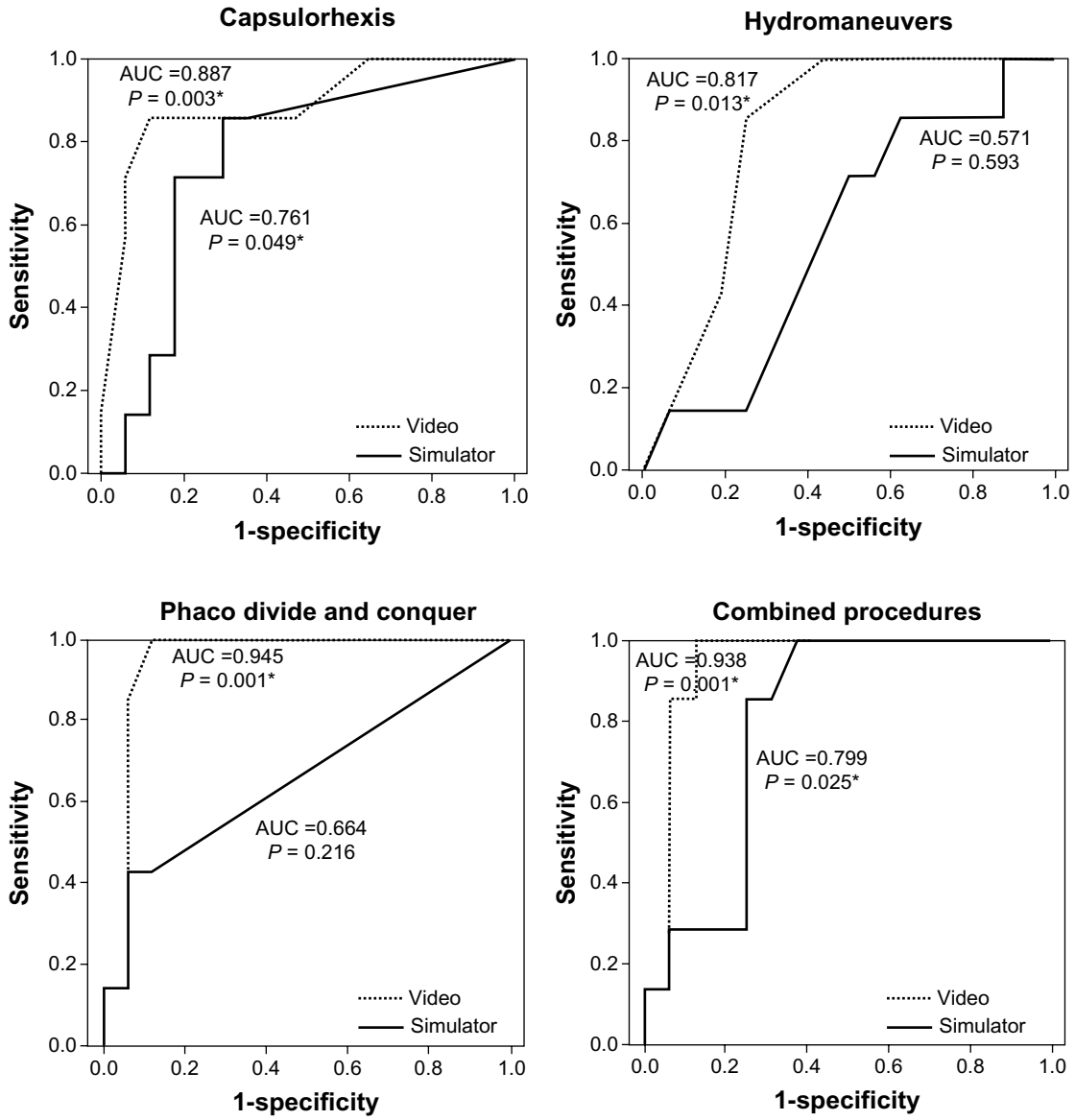

Figure 2 Receiver operating-characteristic curves representing discrimination of cataract surgeons from naïve individuals for capsulorhexis, hydromaneuvers, and phacoemulsification (phaco) divide-and-conquer modules, as well as combined procedures. Area under the curve (AUC) values are shown for simulator score and video score (Objective Structured Assessment of Cataract Surgical Skill for capsulorhexis, hydromaneuvers, phaco divide and conquer, Objective Structured Assessment of Surgical Skills for combined procedure), with respective *P*-values representing difference from hazard area (0.5). *Statistical significance.

video scoring for capsulorhexis (ROC areas 0.835 and 0.965, respectively) and simulator scoring for hydromaneuvers (ROC area 0.581). On the other hand, it did not improve the ROC curve for the phaco-module simulator scoring (ROC area 0.653), further suggesting that the scoring needs to be developed further.

Our study was done with study participants' first encounter with the simulator. One could argue that with a few more training iterations, the discrimination between surgeons and naïve operators would be easier. However, for an ideal simulator, surgeons should score well even without more extensive simulator practice. Furthermore, since the surgeons' activities yielded good discrimination based on the videos, the problem seems to be related to the scoring levels rather than the simulator as such. Another issue with the innate scoring is the truncation of "poor performance," rendering zero points at sometimes minor mistakes. Also, the innate simulator score measures how well the goal was reached, how long the performance took, and the number of actual adverse events. It seems that the human-based scoring also captured potentially hazardous situations that the trainee encountered, besides actual performance. It would be beneficial if the simulator-scoring protocol would also address these cognitive processes.

The ability to distinguish cataract surgeons from naïve individuals shown in this study supports the construct validity of the Eyesi surgical simulator. Cataract surgeons in training can train with the simulator, and with training acquire skills that experienced surgeons have. It is known that complication rates, such as for posterior capsular rupture, are much higher for new cataract surgeons compared to experienced ones, and that it takes about 160 operations before these learning curves associated with cataract training flatten out.[17] Our belief is that the Eyesi intraocular surgery simulator is a beneficial training tool for aspiring cataract surgeons. For it to be used as an assessment tool based on the simulator scoring requires refinement of scoring though. Previous studies have shown concurrent validity for the capsulorhexis procedure, but our suggestion is that, at least for the phaco divide-and-conquer and hydromaneuvers modules, it is more reliable to evaluate the video-based procedure if the simulator is to be used as an assessment tool.[9,12] Further development of the computer-based scoring system, taking into account the cognitive processes embedded in the human-based evaluation, is likely to be a successful route that deserves further research.

Acknowledgments

Data were presented in part at the meeting for the Association for Research in Vision and Ophthalmology, Fort Lauderdale, FL, USA in May 2012, and are part of a PhD thesis: Validity and Usability of a Virtual Reality Intraocular Surgical Simulator. This study was supported by a grant from the Herman Järnhardt Foundation, Malmö, Sweden, and a grant from Lund University. The authors thank Associate Professor Pär-Ola Bendahl, Lund University, for providing the software for the statistical analysis when comparing ROC curves.

Disclosure

The authors have no commercial or proprietary interest in the instrument described.

References

1. Ahmed Y, Scott IU, Greenberg PB. A survey of the role of virtual surgery simulators in ophthalmic graduate medical education. *Graefes Arch Clin Exp Ophthalmol*. 2011;249(8):1263–1265.
2. Oetting TA, Lee AG, Beaver HA, et al. Teaching and assessing surgical competency in ophthalmology training programs. *Ophthalmic Surg Lasers Imaging*. 2006;37(5):384–393.
3. Rogers GM, Oetting TA, Lee AG, et al. Impact of a structured surgical curriculum on ophthalmic resident cataract surgery complication rates. *J Cataract Refract Surg*. 2009;35(11):1956–1960.
4. 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.
5. Mahr MA, Hodge DO. Construct validity of anterior segment antitremor and forceps surgical simulator training modules: attending versus resident surgeon performance. *J Cataract Refract Surg*. 2008;34(6):980–985.
6. Feudner EM, Engel C, Neuhann 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.
7. Belyea DA, Brown SE, Rajjoub LZ. Influence of surgery simulator training on ophthalmology resident phacoemulsification performance. *J Cataract Refract Surg*. 2011;37(10):1756–1761.
8. Saleh GM, Gauba V, Mitra A, Litwin AS, Chung AK, Benjamin L. Objective structured assessment of cataract surgical skill. *Arch Ophthalmol*. 2007;125(3):363–366.
9. Selvander M, Asman P. Virtual reality cataract surgery training: learning curves and concurrent validity. *Acta Ophthalmol*. 2012;90(5):412–417.
10. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg*. 2004;91(2):146–150.
11. Ezra DG, Aggarwal R, Michaelides M, et al. Skills acquisition and assessment after a microsurgical skills course for ophthalmology residents. *Ophthalmology*. 2009;116(2):257–262.
12. Selvander M, Asman P. Cataract surgeons outperform medical students in Eyesi virtual reality cataract surgery: evidence for construct validity. *Acta Ophthalmol*. 2013;91(5):469–474.
13. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–845.
14. Tsuda S, Scott D, Doyle J, Jones DB. Surgical skills training and simulation. *Curr Probl Surg*. 2009;46(4):271–370.
15. Brinkman WM, Buzink SN, Alevizos L, de Hingh IH, Jakimowicz JJ. Criterion-based laparoscopic training reduces total training time. *Surg Endosc*. 2012;26(4):1095–1101.
16. Gauger PG, Hauge LS, Andreatta PB, et al. Laparoscopic simulation training with proficiency targets improves practice and performance of novice surgeons. *Am J Surg*. 2010;199(1):72–80.
17. Randleman JB, Wolfe JD, Woodward M, Lynn MJ, Cherwek DH, Srivastava SK. The resident surgeon phacoemulsification learning curve. *Arch Ophthalmol*. 2007;125(9):1215–1219.