

Comparison of video-based observation and direct observation for assessing the operative performance of residents undergoing phacoemulsification training

Leila Ghiasian, Ali Hadavandkhani, Parya Abdolalizadeh, Leila Janani¹, Acieh Es'haghi

Purpose: To compare the video observation of procedural skills (VOPS) method with the direct observation of procedural skills (DOPS) method in the assessment of senior residents' performance utilizing the International Council of Ophthalmology's Ophthalmology Surgical Competency Assessment Rubric for phacoemulsification (ICO-OSCAR; phaco). **Methods:** This is a prospective comparative study conducted at a university-affiliated hospital. Six ophthalmology residents of postgraduate year 4 participated. Their performance in phacoemulsification was rated via DOPS and later in a masked manner through VOPS by a single faculty assessor. **Results:** Seventy-one surgeries were evaluated. There were no statistically significant differences between the scores of VOPS and DOPS regarding all ICO-OSCAR indices except "instrument insertion into the eye" in which DOPS had higher scores ($P = 0.035$). A significant correlation was observed in total scores of "task-specific" ($r = 0.64$, $P < 0.001$) and "global" ($r = 0.38$, $P = 0.003$) indices between VOPS and DOPS while some subscales did not show a correlation between the two methods of assessment. The Bland-Altman analysis demonstrated that nearly all data points of total "task-specific" and "global" scores fell within the 95% limits of agreement ($[-5.84, 6.87]$ and $[-4.78, 4.86]$, respectively). **Conclusion:** This study demonstrated that VOPS holds promise for a general rating of residents' performance.

Key words: Direct observation, ophthalmology, phacoemulsification, residents, video observation

Direct observation of procedural skills (DOPS) is an acceptable workplace-based assessment method for procedural skills, in which the examiner observes the trainee during a routine procedure on a real patient and in a real situation.^[1-3] The incorporation of cameras into the operating room provides video observation of procedural skills (VOPS) in which common surgical procedures are recorded for assessment and training by self, peers, or faculty reviews.

Few studies have investigated the reliability of the VOPS in procedural specialists.^[4-8] Although some^[5,7,9,10] reported VOPS is a feasible and reliable assessment method for procedural skills, others^[4,8] found this approach inferior to DOPS. There is only one study in ophthalmology proving the reliability of VOPS in trabeculectomy.^[9] The VOPS in phacoemulsification, an intraocular surgery may not have similar reliability to trabeculectomy, an extra and intraocular surgery. Therefore, this study aimed to compare VOPS with DOPS in evaluating the skills of senior residents in performing phacoemulsification surgery.

Methods

This is a prospective study conducted from April 2018 to January 2019 at a university-based hospital. The study adhered to the

Eye Research Center, The Five Senses Institute, Rassoul Akram Hospital, Iran University of Medical Sciences, ¹Department of Biostatistics, School of Public Health, Iran University of Medical Sciences, Tehran, Iran

Correspondence to: Dr. Parya Abdolalizadeh, Rassoul Akram Hospital, Sattarkhan Niayesh St., Tehran 1455364, Iran. E-mail: paryaabdolalizadeh@gmail.com

Received: 26-Apr-2020
Accepted: 15-Aug-2020

Revision: 19-Jul-2020
Published: 17-Feb-2021

Access this article online

Website:
www.ijo.in

DOI:
10.4103/ijo.IJO_1166_20

Quick Response Code:



tenets of the Declaration of Helsinki and the ethic committee's approval was obtained. Informed consent was obtained from residents and the patients separately for using their surgical recordings. Six ophthalmology residents of postgraduate year 4 with an experience of 80 to 120 phacoemulsification surgeries participated in the study. The resident performance was rated via DOPS utilizing the International Council of Ophthalmology's Ophthalmology Surgical Competency Assessment Rubric for phacoemulsification (ICO-OSCAR; phaco)^[11] by one faculty member (an anterior segment specialist) to eliminate the interobserver error. The rater directly visualized through the side glasses of the microscope while she did not guide the resident verbally and did not participate in the surgery during the DOPS assessment. All eyes had phacoemulsification with clear cornea 2.8-mm incision under topical or general anesthesia. No sub-tenon or epi-bulbar augmentation was performed. A continuous curvilinear capsulorhexis was made, and hydrodissection and hydrodelineation were performed. The nucleus was removed by phacoemulsification (Infiniti Vision System, Alcon, Fort Worth, TX, USA) using the stop-and-chop technique. An

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

Cite this article as: Ghiasian L, Hadavandkhani A, Abdolalizadeh P, Janani L, Es'haghi A. Comparison of video-based observation and direct observation for assessing the operative performance of residents undergoing phacoemulsification training. Indian J Ophthalmol 2021;69:574-8.

aspheric intraocular lens (AcrySof®IQ Monofocal IOL, Alcon, USA) was implanted in the bag in all cases, and then the viscoelastic material was washed out of the anterior chamber. In cases with intraoperative complications, the surgical steps before the occurrence of complications were included and the rating was stopped when the residents received any help to finish complicated cases. Full phacoemulsification surgeries were recorded through a Zeiss OPMI Lumera T surgical microscope (Carl Zeiss Meditec, Jena, Germany). Recordings did not include audio. The videotapes were sent to an independent person to remove any logos or other characteristics that could identify the surgeon or date and location of surgery. The videotapes were anonymized and randomized. Later, videotapes were reviewed and graded by the same rater based on VOPS and using the same assessment tool (ICO-OSCAR; phaco).

The ICO-OSCAR; phaco assessment tool consists of a 6-item global rating system (eye positioned centrally, tissue handling, intraocular special awareness, iris protection, overall speed) and a 14-item task-specific checklist (draping, incision, viscoelastic, capsulorhexis flap formation, capsulorhexis circular completion, hydrodissection, instrument insertion into the eye, the stability of the instrument, sculpting, nucleus rotation, nucleus cracking, irrigation, lens insertion, wound closure, wound neutrality) breakdown of cataract surgery. Each global and task-specific component is rated as a novice (score 2), beginner (score 3), advanced beginner (score 4), and competent (score 5).^[11] The main outcome measure of the current study was to correlate the global and task-specific scores of DOPS and VOPS.

All the statistical analyses were performed using SPSS for Windows software version 22 (SPSS, Inc, USA). Wilcoxon test and paired *t*-test were used to compare the scores of VOPS and DOPS. The scores between the two methods were also compared using Spearman’s correlation and Bland-Altman analysis.

Results

Seventy-one cataract surgeries with their videotapes were evaluated via direct and video observation. More than 80% of surgeries (83.1%, 54/65) were done before 3 pm. Topical anesthesia was the most common type of anesthesia (90.0%, 63/70). Seven surgeries became complicated due to anterior capsular tear or posterior capsular rupture. Table 1 shows the basic characteristics of the surgeries.

There was no statistically significant difference between mean DOPS and mean VOPS scores regarding ICO-OSCAR indices [Table 2] except “phacoemulsification/probe insertion into the eye” ($P = 0.035$), in which DOPS score was more than VOPS. Furthermore, significant correlations between DOPS and VOPS were found in some rather than all tasks, including “draping,” “viscoelastic,” “capsulorhexis commencement of flap,” “sculpting,” “nucleus cracking,” and “lens insertion” [Table 2]. However, there was a strong correlation between DOPS and VOPS regarding total task scores ($r = 0.64$, $P < 0.0001$). In Bland-Altman analysis of total task scores, nearly all points located within 95% limits of agreement (-5.84, 6.87). The mean difference between DOPS and VOPS was 0.51. The differences of total task scores (DOPS total score – VOPS total score) were not affected significantly by the time of surgery ($P = 0.284$), type of instrument ($P = 0.682$), and type of anesthesia ($P = 0.697$).

All global scores were similar between DOPS and VOPS [Table 3]. A significant correlation between the two methods was observed only in “tissue handling,” “iris protection,” and “overall speed” [Table 3]. In addition, total global scores of the two methods had significant correlation ($r = 0.38$, $P = 0.003$). In Bland-Altman analysis of total global scores, nearly all points were located within 95% limits of agreement (-4.78, 4.86). The mean difference between global scores of DOPS and VOPS was 0.04. Furthermore, differences of total global scores (DOPS total score – VOPS total score) were not affected significantly by the time of surgery ($P = 0.585$), the type of instrument ($P = 0.498$), and the type of anesthesia ($P = 0.287$).

Discussion

DOPS is a traditional method of assessment that relies upon real-time evaluation of the trainee by the examiner and provides valuable educational effects and timely and immediate feedback.^[3,12,13] Live assessment in DOPS also enables all aspects of the surgeries to be considered, as the change of the surgeon, communication (both verbal and nonverbal) of the operating team, and equipment errors during surgery.^[12] Some shortcomings have been reported in DOPS. Assessors cannot be masked to the identity of the trainee in the DOPS method, introducing the risk of observer bias. Furthermore, the presence of the assessor in the operating room is not only time-consuming but also can expose trainees to stress and nervousness, which disproportionately affect their usual performance in DOPS.^[3]

VOPS was introduced later for workplace-based assessment and it has some advantages over DOPS. One of the most important benefits of VOPS is that it alleviates the time-consuming nature and coordination of the assessor’s schedules to be present in the operating room. Raters can assess whenever having adequate time. Thus VOPS method not only increases efficiency but also limits raters’ burnout from fatigue and loss of concentration. Adequate time and a proper setting for precise, step-by-step assessment of the procedure allow more meticulous detection of any errors in VOPS, which causes lower scores in comparison with DOPS.^[9] This method also offers anonymity, which increases assessment objectivity and eliminates observational bias not to mention the fact that it reduces the anxiety of the trainees during the assessment.^[5]

Table 1: Basic characteristics of 71 phacoemulsification surgeries

Variable	Subgroups	Frequency
Time of Surgery	Early Morning (8-11 am)	38.5% (25/65)
	Mid-Day (11-15)	44.6% (29/65)
	Afternoon (15-19)	16.9% (11/65)
Instrument	Infiniti	61.5% (40/65)
	Stellaris	38.5% (25/65)
Anesthesia	Topical	90.0% (63/70)
	General	10.0% (7/70)
Interruption of Surgery	Anterior Capsular Tear	4.29% (3/70)
	Posterior Capsular Rupture	4.29% (3/70)
	Others	1.43% (1/70)

Table 2: Task-specific scores of 71 phacoemulsification surgeries

ICO-OSCAR Indices	DOPS Mean (SD) Median (IQR), (min, max)	VOPS Mean (SD) Median (IQR), (min, max)	P	Correlation r (P)
Draping	4.41 (0.71) 5 (4,5), (3,5)	4.33 (0.72) 4 (4,5), (3,5)	0.367 [†]	0.46 (P<0.001) [‡]
Incision and Paracentesis	4.61 (0.60) 5 (4,5), (3,5)	4.57 (0.59) 5 (4,5), (3,5)	0.602 [†]	0.25 (P=0.053) [‡]
Viscoelastic	4.70 (0.52) 5 (4,5), (3,5)	4.59 (0.62) 5 (4,5), (3,5)	0.225 [†]	0.41 (P<0.001) [‡]
Capsulorhexis Commencement of Flap	4.49 (SD=0.56) 5 (4,5), (3,5)	4.35 (SD=0.69) 4 (4,5), (3,5)	0.181 [†]	0.27 (P=0.034) [‡]
Capsulorhexis Circular Completion	4.57 (SD=0.65) 5 (4,5), (3,5)	4.56 (SD=0.60) 5 (4,5), (3,5)	0.276 [†]	0.21 (P=0.121) [‡]
Hydrodissection	4.58 (SD=0.63) 5 (4,5), (3,5)	4.59 (SD=0.57) 5 (4,5), (3,5)	0.705 [†]	0.22 (P=0.099) [‡]
Phacoemulsification/Probe and Second Instrument Insertion Into Eye	4.89 (SD=0.36) 5 (5,5), (3,5)	4.80 (SD=0.40) 5 (5,5), (4,5)	0.035 [†]	0.21 (P=0.139) [‡]
Phacoemulsification/ Effective Use and Stability of Instrument	4.48 (SD=0.59) 5 (4,5), (3,5)	4.49 (SD=0.58) 5 (4,5), (3,5)	0.695 [†]	0.20 (P=0.145) [‡]
Sculpting or Primary Chop of nucleus	4.34 (SD=0.71) 4 (4,5), (3,5)	4.20 (SD=0.67) 4 (4,5), (3,5)	0.172 [†]	0.28 (P=0.040) [‡]
Nucleus Rotation and Manipulation	4.14 (SD=0.52) 4 (4,4), (3,5)	4.34 (SD=0.55) 4 (4,4), (3,5)	0.061 [†]	0.27 (P=0.052) [‡]
Nucleus Cracking or Chopping With Safe Phacoemulsification of Segments	3.97 (SD=0.60) 4 (4,4), (3,5)	3.79 (SD=0.65) 4 (3,4), (3,5)	0.144 [†]	0.51 (P<0.001) [‡]
Irrigation and Aspiration	4.22 (SD=0.55) 4 (4,5), (3,5)	4.46 (SD=0.61) 5 (4,5), (3,5)	0.083 [†]	0.26 (P=0.057) [‡]
Lens Insertion	4.48 (SD=0.70) 5 (4,5), (3,5)	4.60 (SD=0.63) 5 (4,5), (3,5)	0.682 [†]	0.34 (P=0.016) [‡]
Wound Closure	4.50 (SD=0.67) 5 (4,5), (3,5)	4.71 (SD=0.54) 5 (4,5), (3,5)	0.260 [†]	0.12 (P=0.394) [‡]
Total Score	62.56 (SD=4.00) 63 (60,65), (51,70)	62.05 (SD=3.90) 62 (58,66), (56,68)	0.330*	0.64 (P<0.001) [§]

*Paired t-test, [†]Wilcoxon test, [‡]Spearman's Correlation Coefficient, [§]Pearson Correlation Coefficient, IQR: Interquartile Range, SD: Standard Deviation, VOPS: Video Observation of Procedural skills, DOPS: Direct observation of Procedural Skills

Table 3: Global Scores of 71 phacoemulsification surgeries

ICO-OSCAR Indices	DOPS Mean (SD) Median (IQR), (min, max)	VOPS Mean (SD) Median (IQR), (min, max)	P	Correlation r (P)
Wound Neutrality and Minimizing Eye Rolling and Corneal Distortion	4.25 (SD=0.60) 4 (4,5), (3,5)	4.17 (SD=0.67) 4 (4,5), (3,5)	0.707 [†]	0.24 (P=0.090) [‡]
Eye Positioned Centrally Within Microscope View	4.14 (SD=0.69) 4 (4,5), (3,5)	3.98 (SD=0.59) 4 (4,4), (3,5)	0.364 [†]	0.02 (P=0.910) [‡]
Conjunctival and Corneal Tissue Handling	4.60 (SD=0.71) 5 (4,5), (3,5)	4.58 (SD=0.56) 5 (4,5), (3,5)	0.686 [†]	0.43 (P=0.002) [‡]
Intraocular Spatial Awareness	4.71 (SD=0.50) 5 (4,5), (3,5)	4.75 (SD=0.53) 5 (4,5), (3,5)	0.201 [†]	0.23 (P=0.105) [‡]
Iris Protection	4.75 (SD=0.54) 5 (5,5), (3,5)	4.84 (SD=0.46) 5 (5,5), (3,5)	0.480 [†]	0.60 (P<0.001) [‡]
Overall Speed and Fluidity of Procedure	4.39 (SD=0.58) 4 (4,5), (3,5)	4.45 (SD=0.57) 4 (4,5), (3,5)	0.655 [†]	0.50 (P<0.001) [‡]
Total Global Scores	26.71 (SD=2.40) 27 (26,28), (21,30)	26.83 (SD=2.19) 27 (25,29), (21,30)	0.909*	0.38 (P=0.003) [§]

*Paired t-test, [†]Wilcoxon test, [‡]Spearman's Correlation Coefficient, [§]Pearson Correlation Coefficient, IQR: Interquartile Range, SD: Standard Deviation

The VOPS method has some drawbacks. First, technical problems make some records unavailable for assessment via VOPS and inadequate technological support in hospitals may

be a barrier to its utilization.^[4,8] The utility of video remains dependent on optimal views and standardized editing. There was a decentered view of the camera in VOPS despite the

centered microscope view during DOPS. This discrepancy led to no significant correlation between the scores of DOPS and VOPS methods in the index of *eye positioned centrally* in the current study. Poor correlation between the two methods in some indices, such as *intraocular special awareness*, *wound neutrality and corneal distortion*, and *hydrodissection* in our study may be attributed to the lower quality of vision of the operative field in the video recordings in comparison to direct observation via surgical microscope.^[14] On the other hand, reaching an optimal view requires a higher resolution of video recording in intraocular surgeries, such as phacoemulsification due to high magnification and fine movements. This may explain the weaker correlation between DOPS and VOPS methods in the current study in comparison with Hassanpour's study^[9] on trabeculectomy as an extraocular surgery.

Secondly, some crucial information is lost in the VOPS method.^[4] The records contain visual information from the microscope, while neither audio information nor visual information from the external operating room environment is recorded. This may not provide enough information for the rater to assess some aspects of the performance leading to misevaluation.^[4,9] For example, the evaluator does not detect the equipments' ineffectiveness, especially the phaco machine, and can underrate the performance of the resident in the VOPS method compared to the DOPS method. Although the current study did not show any significant difference between scores among nearly all subscales of ICO-OSCAR based on DOPS and VOPS, the absence of correlation between these two methods in some task-specific indices, such as *effective use and stability of phaco probe* and *irrigation and aspiration* may be due to this limitation of the VOPS method. Scott *et al.*^[4] also reported a poor correlation between the scores of VOPS and DOPS in the evaluation of laparoscopy. They mentioned that the VOPS method did not provide enough information about "knowledge of instruments," "use of assistants," and "knowledge of specific procedure."^[4] Hassanpour *et al.*^[9] also observed the scores of VOPS are lower than DOPS in the evaluation of residents' skills in trabeculectomy despite the significant correlation between these two methods.

Common surgical evaluations involve a single global rating made by the supervisors.^[2,15] Global scores rate the residents' skills and the adequacy of his or her technical proficiency using general performance as a benchmark and are, therefore, more widely applicable and demonstrate superior reliability and validity, although it cannot be considered an adequate assessment of technical skill.^[5,16,17] In contrast, task-specific scores are procedure-specific and require observation of the entire surgery. They provide a greater degree of formative feedback to the trainee through the identification of areas of weakness.^[12,14] We used a validated and objective performance-rating tool, (the ICO-OSCAR; phaco questionnaire) consisting of standardized criteria with *global* and *task-specific* components, each rated on a five-point Likert scale.^[17] This tool provides specific guidelines for grading each surgical step.^[11] Current study demonstrated significant correlation between the VOPS and DOPS methods regarding the *total* scores of task-specific indices ($r = 0.64$, $P < 0.001$) and the *total* scores of global indices ($r = 0.38$, $P = 0.003$). An acceptable agreement between the two methods was also observed in the Bland-Altman plot using *total* task-specific and global scores. Furthermore, the VOPS method showed significant correlation with the DOPS method in some global ratings including *overall speed and*

fluidity of procedure ($r = 0.50$, $P < 0.001$), *iris protection* ($r = 0.60$, $P < 0.001$), and *tissue handling* ($r = 0.43$, $P = 0.002$). In contrast, the correlation between VOPS and DOPS was neither statistically nor weakly significant ($r \leq 0.3$) in 11 out of 14 task-specific subscales [Table 1]. Therefore, our study suggests that the VOPS method can be replaced by the DOPS assessment in clinical practice when we aim to evaluate only the general performance of the residents rather than detailed steps of the phacoemulsification. Hassanpor *et al.*^[9] also recommended the VOPS as an alternative method for the evaluation of residents' skills in trabeculectomy. However, they did not evaluate *mini skills* of trabeculectomy separately and just used the total scores.^[9]

The present study has some limitations. This study was done only in one university-based hospital with small sample size. Recruiting a single assessor is another limitation of the study. Although anonymization and randomization were done to prevent recall bias in VOPS, it might not be sufficient to erase the memory of the rater who participated in DOPS. Moreover, the current study did not evaluate different operating microscopes with different resolutions and image quality which may affect the VOPS. Further studies with more sample size, with at least two assessors and different operating microscopes, are needed to verify the results.

Conclusion

In conclusion, this is the first study that compares the DOPS and VOPS methods for evaluation of the skills of senior residents in performing phacoemulsification surgery, using ICO-OSCAR; phaco questionnaire. The results show that the method of video-based observation can be adapted in the general evaluation process of residents' performance in phacoemulsification while it may not be equal to the method of direct observation regarding the details.

Financial support and sponsorship

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

There are no conflicts of interest.

References

1. Brown N, Doshi M. Assessing professional and clinical competence: The way forward. *Adv Psychiatr Treat* 2006;12:81-9.
2. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach* 2007;29:855-71.
3. Fromme HB, Karani R, Downing SM. Direct observation in medical education: A review of the literature and evidence for validity. *Mt Sinai J Med* 2009;76:365-71.
4. Scott DJ, Rege RV, Bergen PC, Guo WA, Laycock R, Tesfay ST, *et al.* Measuring operative performance after laparoscopic skills training: Edited videotape versus direct observation. *J Laparoendosc Adv Surg Tech* 2000;10:183-90.
5. Aggarwal R, Grantcharov T, Moorthy K, Milland T, Darzi A. Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Ann Surg* 2008;247:372-9.
6. Barton JR, Corbett S, Van der Vleuten CP. The validity and reliability of a Direct Observation of Procedural Skills assessment tool: Assessing colonoscopic skills of senior endoscopists. *Gastrointest Endosc* 2012;75:591-7.

7. Jain NS, Schwarzkopf R, Scolaro JA. Video review as a tool to improve orthopedic residents' performance of closed manipulative reductions. *J Surg Educ* 2017;74:663-7.
8. Beckmann RB, Lipscomb GH, Ling FW, Beckmann CA, Johnson H, Barton L. Computer-assisted video evaluation of surgical skills. *Obstet Gynecol* 1995;85:1039-41.
9. Hassanpour N, Chen R, Baikpour M, Moghimi S. Video observation of procedural skills for assessment of trabeculectomy performed by residents. *J Curr Ophthalmol* 2016;28:61-4.
10. Driscoll PJ, Paisley AM, Paterson-Brown S. Video assessment of basic surgical trainees' operative skills. *Am J Surg* 2008;196:265-72.
11. Golnik KC, Beaver H, Gauba V, Lee AG, Mayorga E, Palis G, *et al.* Development of a new valid, reliable, and internationally applicable assessment tool of residents' competence in ophthalmic surgery (An American Ophthalmological Society Thesis). *Trans Am Ophthalmol* 2013;111:24-33.
12. Cremers SL, Ciolino JB, Ferrufino-Ponce ZK, Henderson BA. Objective assessment of skills in intraocular surgery (OASIS). *Ophthalmology* 2005;112:1236-41.
13. Wilkinson JR, Crossley JGM, Wragg A, Mills P, Cowan G, Wade W. Implementing workplace based assessment across the medical specialties in the United Kingdom. *Med Educ* 2008;42:364-73.
14. Bhogal MM, Angunawela RI, Little BC. Use of low-cost video recording device in reflective practice in cataract surgery. *J Cataract Refract Surg* 2010;36:542-6.
15. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, *et al.* Objective structured assessment of technical skills (OSATS) for surgical residents. *Br J Surg* 1997;84:273-8.
16. Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med* 1998;73:993-7.
17. Beard JD, Choksy S, Khan S. Assessment of operative competence during carotid endarterectomy. *Br J Surg* 2007;94:726-30.