Exploring the Assessment Tool ICO OSCAR. A Comparison of Reliability of Whole of the Operation with Its Parts

Khawaja Khalid Shoaib¹, Judy Scopes², Zunaira Mubarik³, Fiza Azhar⁴, Amna Mehmud⁵
¹,³,⁴,⁵Mughal Eye Hospital, Lahore, ²Centre for Medical Education, Dundee University, UK

ABSTRACT

Purpose: To compare the reliability of Ophthalmology Surgical Competency Assessment Rubric (ICO-OSCAR: phaco) between assessment of whole of the operation and its parts.

Study Design: Mixed method study design.

Place and Duration of Study: Mughal Eye Hospital, Lahore from January 2020 to January 2021.

Methods: Three raters assessed six anonymized phaco video operations with the help of ICO OSCAR phaco first as whole operations and then as parts of the operation. Inter rater reliability of the assessments of parts of the operation were compared with inter rater reliability of whole of the operation. Intra class coefficient (ICC) was used in SPSS – version 20. Raters also filled a survey form to detect how satisfied they were with the rubric. Raters’ feedback was taken to describe shortcomings/faults found during the use of ICO OSCAR phaco form.

Results: In the quantitative part, the ICC for the combined parts of the operation (0.910) was better than the value of ICC for the whole operation (0.904). In the questionnaire part, raters were satisfied with ICO OSCAR phaco rubric as a useful tool for learning and assessment of phacoemulsification surgical skills. In the qualitative part, many deficiencies were observed by the assessors in the rubric during videos analysis with the rubric ICO OSCAR phaco. This rubric is more accurate for assessment of operations in the operation theater.

Conclusion: Parts of the operations can be used instead of the whole operations for the assessment of surgical skills of phacoemulsification with the help of ICO OSCAR phaco rubric. For assessment of videos, it requires some modifications as it is more valid for the assessment of surgical skills during live surgery than in operation theater.

Key Words: ICO OSCAR phaco rubric, phacoemulsification, Cataract.


Doi: 10.36351/pjo.v39i2.1545

INTRODUCTION

Cataract is a major cause of treatable blindness. Phacoemulsification (phaco) is the most commonly performed surgery for cataract and hence most commonly taught in all surgical subspecialty institutions. Supervised surgery on animal eyes, simulators, wet lab, videos of operations and discussion with senior surgeons have been studied for their effectiveness in learning surgical skills.¹ Assessment drives learning and to ensure that learning has taken place, teaching should be supplemented with assessment.² A good assessment tool must have certain characteristics which include validity, reliability, feasibility and acceptability.

“Reliability relates to the consistency of a measure”³. In other words, “consistency of a test over time, over different cases (inter-case), and different
explorers (inter-rater). Inter rater reliability measures the consistency of rating of performance by different examiners (raters) keeping all the other variables as consistent as possible.\textsuperscript{4} If two or more assessors or raters give similar grade to a video of operation and correlation analysis confirms the similarity, it is called good or high inter-rater reliability or low variability.

Validity “determines whether an assessment instrument really tests what it is supposed to test”.\textsuperscript{4} In other words “Validity is defined as “the extent to which a concept is accurately measured in a quantitative study”.”\textsuperscript{3}

Simulators have been found to be comparable to wet lab learning for strabismus surgery.\textsuperscript{6} Eyesi direct and indirect ophthalmoscopes have also been investigated though less extensively.\textsuperscript{7} Literature shows that training of junior surgeons on Eyesi simulator decreases cataract surgery complication rates.\textsuperscript{8,9} Another way to train junior surgeons is with video recording. Videos of cataract surgery can be augmented with instrument labels to automatically detect phases and formation of algorithms. However, there is still requirement of advancement of technology as far as its role in education and to make it comparable to human evaluation is concerned.\textsuperscript{9}

A tool can be used without hesitation in any assessment if its reliability has been confirmed. These tools can be used for self-assessment by trainees. Virtual reality simulation based\textsuperscript{10} and wet lab training are now most important resources of learning cataract surgery.\textsuperscript{11,12} However, surgical videos and rubrics scoring are superior to scoring based on simulators.\textsuperscript{13}

One of the scoring tools is Ophthalmology Surgical Competency Assessment Rubric (OSCAR) was popularized by International Council of Ophthalmology (ICO) and its version for phaco is the ICO OSCAR: phaco. We used this scoring tool to compare between the reliability of assessment of whole surgical procedure and assessment of its different parts.

METHODS
This is a mixed method study conducted in Mughal eye hospital Lahore. In sequential pattern first quantitative and then qualitative data was collected. In the first part, inter-rater reliability of parts of operations was compared with inter-rater reliability of whole operations and a questionnaire was presented to raters to describe how useful the rubric ICO OSCAR Phaco was. Second part was qualitative in which assessors/raters were asked to describe deficiencies of the rubric and the problems they encountered during marking of the videos with the rubric.

Six phacoemulsification procedures were video recorded. Inclusion Criteria for surgical procedures were uncomplicated, full cases of phaco which had no specific recognizable markings on the screen when displayed. Exclusion criteria comprised of any intervention that was required in addition to routine phaco operation e.g., requirement for additional anesthesia (peribulbar/sub tenon etc.), posterior capsular rupture with/without vitreous loss etc. and prolonged procedures.

The recordings were done by operation theatre assistants/junior doctors who did not participate in the assessment part of the research project. The videos contained one or more operations of individual surgeons. These surgeons/trainees were selected on a voluntary basis and they had different levels of experience. This ranged from junior to senior surgeon. The videos were anonymized and assessed by three assessors (initially four assessors were selected but one dropped out because of personal commitments and could not complete the assessment). The assessment was done in the light of the rubric ICO OSCAR: phaco.

The assessors were selected from a local eye hospital based on their interest in research. They had varying experience in phaco operations. They did the job of assessing videos on voluntary basis. First assessor had experience of 700 phaco operations. However, this rater dropped out and his assessment data was excluded from the study. His opinion on the feasibility of the rubric was included. Second assessor (C) had experience of 200 phaco operations, third assessor had less experience of only 2 phaco operations but had assisted dozens of phaco (B). The fourth assessor, the most junior, had no experience of doing independent phaco operations but assisted a lot of phaco operations (A).

A laptop was used to show video clips to the assessors. Same magnification and illumination settings were used by all the assessors to prevent variation in marking due to these parameters. For assessment, the videos were split into four parts with the help of Windows movie maker. Assessors watched the videos and marking was done on the rubric. Marking on paper was followed by its electronic
conversion and saved in the form of tables in Microsoft Word files. ICO Phaco rubric was followed for marking except its No. 1 step “Draping” which was omitted as in these cases draping had been done by the operation theater assistants and not the surgeons/trainees.

In first set of videos, operation number 1 and 2 were shown as full operations and operation No 3, 4 and 5 were divided into four parts. These four parts consisted of:
1st part: From incision to entry of phaco probe.
2nd part: From entry of phaco probe for phaco to taking out of the phaco tip after completing phaco.
3rd part: From start to end of aspiration of cortical material with the help of irrigation aspiration cannula.
4th part: From IOL implantation till end of the operation.

In ICO OSCAR phaco has divided the phaco operation into sixteen steps. Operation was divided into four parts. Thus, each part consisted of four steps.

Videos were divided into two sets so that it was easy to remember for the assessors how many videos they have marked. Each assessor marked both sets of videos.

Table 1: Relationship of videos with operations - Set 1 (video 1 to 6), Set 2 (video 7 to 14).

<table>
<thead>
<tr>
<th>Set 1</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video 1</td>
<td>1st complete operation</td>
</tr>
<tr>
<td>Video 2</td>
<td>2nd complete operation</td>
</tr>
<tr>
<td>Video 3</td>
<td>First parts of operation No 3, 4 and 5</td>
</tr>
<tr>
<td>Video 4</td>
<td>Second parts of operation No. 3, 4 and 5</td>
</tr>
<tr>
<td>Video 5</td>
<td>Third “ “ “</td>
</tr>
<tr>
<td>Video 6</td>
<td>Fourth “ “ “</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Set 2</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video 7</td>
<td>1st parts of operation No. 1, 2 &amp; 6</td>
</tr>
<tr>
<td>Video 8</td>
<td>2nd parts of operation No. 1, 2 &amp; 6</td>
</tr>
<tr>
<td>Video 9</td>
<td>3rd parts of operation No. 1, 2 &amp; 6</td>
</tr>
<tr>
<td>Video 10</td>
<td>4th parts of operation No. 1, 2 &amp; 6</td>
</tr>
<tr>
<td>Video 11</td>
<td>Complete operation No. 3</td>
</tr>
<tr>
<td>Video 12</td>
<td>Complete operation No. 4</td>
</tr>
<tr>
<td>Video 13</td>
<td>Complete operation No. 5</td>
</tr>
<tr>
<td>Video 14</td>
<td>Complete operation No. 6</td>
</tr>
</tbody>
</table>

In the second set, operation No. 1, 2 and 6 were divided into four parts each as described above.

A questionnaire was devised for assessors to gather their views on usefulness (this relates to the questions asked) of the rubric and whether assessment of the parts was better than the whole. It was sent to first rater to detect any ambiguity or confusion in the questionnaire. When no problem was found in understanding it, it was sent to the rest of the raters. The questionnaire was sent to the assessors online by Google forms and their responses were collected and analyzed. It was voluntary, anonymized and without any monetary benefits. Data was collected from assessors/raters about their experience with the ICO OSCAR rubric.

RESULTS
Three assessors analyzed six anonymized video recordings using ICO OSCAR: phaco rubric.

1, 2, 3….6 are operations. Assessors A, B & C gave marks. Each rater assessed each operation twice. Assessment of whole operation (W) and assessment as parts of the operation (P). Their score was entered in Table 3, 4, 5.

Intra class coefficient (ICC) for the whole operation was 0.904 while ICC for the parts of the operation was 0.910, which means that the by parts operation technique is little better than whole operation technique.

Responses to questionnaire were as follows:
1. “ICO OSCAR Phaco is useful for learning/teaching of Phaco surgical skill”. Four (100%) agreed.
2. “This instrument is useful for assessment of Phaco surgical skill”. One (25%) was neutral, 2 (50%) agreed and 1 (25%) strongly agreed.
3. “This instrument is useful for self-assessment of Phaco surgical skill”. One (25%) agreed and 3 (75%) strongly agreed.
4. “Assessment of parts of operations is better than assessment of the whole operations”. One 1 (25%) disagreed, 2 (50%) agreed, 1 (25%) strongly agreed.

Raters were asked to tell deficiencies/confusions/difficulties encountered during their marking.

A – Confusions/deficiencies in the rubric:
1. “Type of visco”, “requires minimum instructions” and “Phaco power used” were the points not obvious in the videos.
2. Endothelial/capsular touch could not be appreciated with average quality videos.
Table 2: Score given by assessors, A, B & C.

<table>
<thead>
<tr>
<th></th>
<th>1 W/P</th>
<th>2 W/P</th>
<th>3 W/P</th>
<th>4 W/P</th>
<th>5 W/P</th>
<th>6 W/P</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>34/35</td>
<td>25/32</td>
<td>18/28</td>
<td>31/27</td>
<td>17/20</td>
<td>34/33</td>
<td>159/175</td>
</tr>
<tr>
<td>1st portion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd portion</td>
<td>29/35</td>
<td>21/25</td>
<td>18/33</td>
<td>23/33</td>
<td>21/27</td>
<td>34/32</td>
<td>146/185</td>
</tr>
<tr>
<td>3rd portion</td>
<td>12/17</td>
<td>10/18</td>
<td>11/14</td>
<td>11/8</td>
<td>14/10</td>
<td>14/15</td>
<td>72/82</td>
</tr>
<tr>
<td>4th portion</td>
<td>22/15</td>
<td>14/13</td>
<td>15/20</td>
<td>15/14</td>
<td>19/19</td>
<td>98/96</td>
<td></td>
</tr>
<tr>
<td>B w/p</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>Total</td>
</tr>
<tr>
<td>1st portion</td>
<td>33/34</td>
<td>31/34</td>
<td>34/32</td>
<td>34/32</td>
<td>33/35</td>
<td>201/197</td>
<td></td>
</tr>
<tr>
<td>2nd portion</td>
<td>31/35</td>
<td>29/34</td>
<td>34/34</td>
<td>34/34</td>
<td>32/35</td>
<td>194/207</td>
<td></td>
</tr>
<tr>
<td>3rd portion</td>
<td>13/15</td>
<td>11/15</td>
<td>14/15</td>
<td>14/14</td>
<td>13/14</td>
<td>79/88</td>
<td></td>
</tr>
<tr>
<td>4th portion</td>
<td>17/20</td>
<td>17/20</td>
<td>19/19</td>
<td>19/19</td>
<td>19/19</td>
<td>109/117</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1W/P</td>
<td>2W/P</td>
<td>3W/P</td>
<td>4W/P</td>
<td>5W/P</td>
<td>6W/P</td>
<td>Total</td>
</tr>
<tr>
<td>1st portion</td>
<td>30/24</td>
<td>24/23</td>
<td>22/28</td>
<td>26/30</td>
<td>34/35</td>
<td>167/175</td>
<td></td>
</tr>
<tr>
<td>2nd portion</td>
<td>28/19</td>
<td>23/21</td>
<td>26/28</td>
<td>27/28</td>
<td>34/28</td>
<td>167/152</td>
<td></td>
</tr>
<tr>
<td>3rd portion</td>
<td>13/8</td>
<td>10/8</td>
<td>10/12</td>
<td>11/11</td>
<td>14/14</td>
<td>71/64</td>
<td></td>
</tr>
<tr>
<td>4th portion</td>
<td>16/15</td>
<td>13/19</td>
<td>14/16</td>
<td>15/13</td>
<td>17/19</td>
<td>93/102</td>
<td></td>
</tr>
</tbody>
</table>

W = whole operation, P = Part of operation, Operation No. 1, 2,...6

3. In point no. 5, it was mentioned that CCC size depend upon type of IOL which was apparent only in the later part of the operation and could not be commented upon in the early stages. Assessor would have to mark this step later.

4. “Working on foot pedals” in no.9 was obvious to person sitting next to the trainee during the operation but not to one who was assessing the video.

5. “Phaco pieces” done or not – it depends upon the procedure.

B – Problematic points with regard to assessments of parts with the ICO OSCAR phaco:

1. Few points of the rubric were more evident in the subsequent steps of the surgery; for example, in point “incision and paracentesis” under score 5, it was mentioned in the rubric that it should be “self-sealing”. This is evident only after intraocular manipulations in the later steps. Therefore, an assessor during the whole procedure assessment can turn to this point later and mark it but if the assessment parts is done, assessor would not be able to comment on it at this stage. For assessment of parts, this point can be placed at a later stage of the operation.

2. In Hydro dissection (point No. 6), if fluid was injected but no fluid wave was visible and surgeon did not rotate the nucleus, the nuclear mobility/capacity to rotate would be evident in the next steps of phaco only.

3. During the first assessor’s marking it was noticed that in one operation, cutting a portion of the video clip at the end of irrigation/aspiration (part No. 3) where a small amount of lens matter was still there resulted in inclusion in the IOL portion (part No. 4). Naturally, it would lead to deduction of marks in 3rd portion and would lead to difference when parts would be compared with the whole operations. So the video was edited again and final portion of aspiration of last part of cortex was removed from part No. 4 and included in part No. 3. A careful division of parts is thus an important consideration when assessment of parts is the aim.

DISCUSSION

The Intra class coefficient for the parts of the operation (0.910) was better than the value of Intra class coefficient for the whole operation (0.904). We can assume that the parts can be used instead of the whole operations for the assessment purposes. In our study raters were untrained. Untrained assessors were selected to mimic the ground reality because at present most of ophthalmology teachers are not trained in the use of ICO OSCAR rubric. Rater training may increase reliability.

As MCQs have a better reliability than essay questions, similarly, video clips having one part of the whole operation can be used to assess trainees/surgeons. It decreases the work required by the qualified persons. It increases the work of Information technology (IT) persons as editing of operation videos is required. However, this job can be done by junior staff, decreasing burden/workload of senior consultants/teachers. Thus, multiple clips of parts of
different operations can be assessed at the same time within the time period taken for assessing only one lengthy operation. It will increase breadth of assessment in terms of number of different types of operations of a single candidate. Assigning different assessors to assessment of different parts of operations will reduce the subjectivity, present in the assessment of operations. Assessment of initial parts may induce bias towards marking of later parts when the whole operations are checked in one go, which is avoided if only selected portions are checked.

A recent review article has summarized the tools used in assessment of surgical skills in Ophthalmology. Different types of operation assessments include: procedure based (checklists/global rating scales), simulation based (dry lab, wet lab and virtual), knowledge based, outcome data and motion analysis.

One may argue that surgical time can be taken as an indicator for competency because surgical time is usually short in expert hands and more in novice. However, it cannot be made the only criteria for gauging competency. Time alone cannot differentiate between expert and novice. It is an important aspect, so it has been included in the global indices of the rubric. Rating was obtained from one surgeon only and case complexity was also not considered. Literature shows that different parts of surgery (for example capsulorhexis) in the ICO OSCAR have been studied for deep learning techniques validation.

Assessment of the whole operation may induce fatigue by considering all options of grading for each step. Secondly, it may lead to bias for later steps marking if earlier steps are different in performance. One step of operation i.e., continuous curvilinear capsulorrhexis (CCC) was studied for inter and intra observer reliability. Twelve questions regarding CCC of surgeons of different experiences were investigated on videos by 7 reviewers for inter observer reliability. Intra-observer reliability was checked after 4 weeks reassessment. It revealed one observer having large variations, one intermediate and the rest 5 minimal variations. One interesting article concluded that surgical trainees were not very accurate in diagnosing their own surgical skill but could identify the surgical performance of their peer adequately.

The rubric was not being used frequently 5 to 10 years previously, even in developed countries. Though the Royal College of Surgeons and Physicians Canada has recommended that residency training should be competency based, a 2017 survey revealed that no institution is using published assessment tools for cataract surgery. For augmenting surgical learning and teaching, wet lab was being used in all and simulators by 45% of the institutions.

In the last 2 – 3 years, situation regarding use of ICO OSCAR phaco has changed and now it is being used in developing countries as well. A recent article from India has highlighted the importance of using rubric ICO OSCAR phaco. A similar article from Pakistan analyzes the use of the rubric. Their conclusion is again same that the rubric is better than traditional methods of teaching phaco.

Institutions around the globe are now offering to provide learning opportunities for phaco operation with their assessment by the ICO OSCAR phaco rubric. Similarly wet lab training through Cyber sight, Orbis international telemedicine platform is offering distance learning and assessment with the help of ICO OSCAR phaco rubric. A consensus is building that learning cataract surgery in a structured training program and proper assessment with ICO OSCAR rubric not only facilitates early conversion from novice to expert but also decreases the complication rate of the operations. Decreased complication rate (of post capsular rent and errant capsulorrhexis) has been confirmed with the use of this rubric.

As per the recommendation of ICO OSCAR phaco, 30 minutes of the phaco operation was used to check for competency based on time duration. Now assessment is being done using this gauge to determine competency. In a recent article, for the assessment of procedural skills, video observation has been compared with direct observation using ICO OSCAR phaco and scores with both were within 95% limits of agreement.

Regarding the questionnaire, assessors overall agreed that the rubric was useful. Limitations of the study was a small sample size. The assessors were also not trained and problems related with the cutting of video into different parts.

If training has to be switched from completing logbooks and compiling lists of procedures done to competency-based training, meeting the demands of learning outcomes, we will have to adopt standardized tools for assessment. It would be better to have videos and photographs hyperlinked in the rubric squares.
CONCLUSION
Despite some deficiencies and ambiguities, ICO OSCAR Phaco is reliable and feasible instrument in assessing parts of operations better than the whole operations.

Conflict of Interest
Authors declared no conflict of interest.

Ethical Approval
The study was approved by the Institutional review board/Ethical review board (AMC/PGMILGH/00-112-20).

REFERENCES


Authors’ Designation and Contribution
Khawaja Khalid Shoaib; Consultant Ophthalmologist: Concepts, Design, Literature search, Data acquisition, Data analysis, Statistical analysis, Manuscript preparation, Manuscript editing, Manuscript review.

Judy Scopes; Supervisor: Design, Data analysis, Manuscript preparation, Manuscript editing, Manuscript review.

Zunaira Mubarik; Consultant Ophthalmologist: Design, Data acquisition, Manuscript preparation, Manuscript review.

Fiza Azhar; Consultant Ophthalmologist: Design, Data acquisition, Manuscript preparation, Manuscript review.

Amna Mehmud; Consultant Ophthalmologist: Design, Data acquisition, Manuscript preparation, Manuscript review.

...★...