HOSTED BY

ELSEVIER

CrossMark

Available online at www.sciencedirect.com

ScienceDirect

Journal of Current Ophthalmology 28 (2016) 61–64

http://www.journals.elsevier.com/journal-of-current-ophthalmology

Original research

Video observation of procedural skills for assessment of trabeculectomy performed by residents

Narges Hassanpour [a], Rebecca Chen [b,c], Masoud Baikpour [d], Sasan Moghimi [a,b,*]

[a] *Eye Research Center, Farabi Eye Hospital, Tehran University of Medical Sciences, Tehran, Iran*
[b] *Koret Eye Research Center, University of California at San Francisco, USA*
[c] *Case Western Reserve University School of Medicine, Cleveland, OH, USA*
[d] *Department of Medicine, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran*

Received 1 February 2016; revised 19 March 2016; accepted 20 March 2016
Available online 8 April 2016

Abstract

Purpose: The efficacy and sufficiency of a healthcare system is directly related to the knowledge and skills of graduates working in the system. In this regard, many different assessment methods have been proposed to evaluate various skills of the learners. Video Observation of Procedural Skills (VOPS) is one newly-proposed method. In this study we aimed to compare the results of the VOPS method with the more commonly used Direct Observation of Procedural Skills (DOPS).

Methods: In this prospective study conducted in 2012, all 10 ophthalmology residents of post graduate year 4 were selected for participation. Three months into training in the glaucoma ward, these residents performed trabeculectomy surgery on patients, and their procedural skills were assessed in real time by an expert via the DOPS method. All surgeries were also recorded and later evaluated via the VOPS method by an expert. Bland–Altman plot also was used to compare the two methods and calculating the mean and 95% limit of agreement.

Results: Residents have been done a mean of 14.9 ± 3.5 (range 10–20) independent trabeculectomy before the assessments. DOPS grade was positively associated with number of independent trabeculectomy during glaucoma rotation (β=0.227, $p = 0.004$). The intra-observer reproducibility of VOPS measurements was 0.847 (95% CI: 0.634, 0.961). The mean VOPS grade was significantly lower than the mean DOPS grade (8.4 vs. 8.9, $p = 0.02$). However, a good correlation was observed between the grades of VOPS and DOPS ($r = 0.89$, $p = 0.001$). Bland–Altman analysis demonstrated that all data points fell within the 95% limits of agreement (−1.46, 0.46).

Conclusion: The present study showed that VOPS might be considered a feasible, valid, and reliable assessment method for procedural skills of medical students and residents that can be used as an alternative to the DOPS method. However, VOPS might underestimate DOPS in evaluating surgical skills of residents.

Keywords: Video observation of procedural skills; Direct observation of procedural skills; Ophthalmology; Residents

Introduction

Medical education has progressed significantly and is going through major changes all around the world.[1] In response to different challenges from society, patients, students, and physicians, medical schools are developing new methods of teaching and assessment.[2,3] The efficacy and sufficiency of a healthcare system is directly related to the skills and abilities of graduates working in the system, which includes not only knowledge and technical skills but also analytical abilities and communication skills.[4] Accordingly, assessment systems must be comprehensive, logical and precise enough to be able to evaluate the required attributes along with assessment of necessary knowledge and skills.[5,6]

* Corresponding author. Farabi Eye Hospital, Eye Research Center, Qazvin SQ, Tehran, 13336, Iran.
E-mail address: sasanimii@yahoo.com (S. Moghimi).
Peer review under responsibility of the Iranian Society of Ophthalmology.

http://dx.doi.org/10.1016/j.joco.2016.03.003

Many different assessment methods have been proposed to evaluate various skills of the trainees, particularly procedural skills. Direct Observation of Procedural Skills (DOPS) is the most commonly used assessment method for evaluating these abilities.[7] In this method, an advanced trainee performs a procedure on a patient, and an experienced and knowledgeable assessor observes the trainee's performance.[7] This method provides high face validity, and the setting of the test closely resembles clinical practice.[8] The most important shortcoming of this approach is that multiple trainees cannot be accurately assessed at the same time. Due to the growing number of residents, the DOPS method will require significant time investment.[9] Another problem is that since residents are aware they are being observed, they might not perform the procedure as they usually do.[9]

Another assessment method recently proposed is Video Observation of Procedural Skills (VOPS), in which the trainee's performance is recorded and later evaluated by the assessor.[10] This method provides the assessor with adequate time and a proper setting for precise step by step assessment of the procedure. This method is more systematic and can be blinded. It also allows the residents to participate in the assessment process and compare their procedural skills with other trainees. The problem with this method is its lower face validity compared to the DOPS method.[10]

One of the important procedures that ophthalmology residents should be able to perform at the end of their training period is trabeculectomy.[11] Since the VOPS method has been proposed recently, few studies are present that have evaluated it validity and feasibility. Accordingly, we aimed to compare the results of VOPS and DOTS assessment methods in evaluating the procedural skills of ophthalmology residents in performing trabeculectomy.

Methods

In this prospective study conducted in 2012, all 10 ophthalmology residents of post graduate year 4 were selected for participation. After three months training in the glaucoma ward of Farabi Eye Hospital and performing trabeculectomy surgery on patients, their procedural skills were evaluated. Trabeculectomy procedure was divided into seven miniskills including: a) peritomy, b) conjunctival dissection, c) Mitomycin application, d) scleal flap creation, e) sclerotomy, f) closing sclera flap with releasable suture, and e) closing conjunctiva and resident's procedural skills were assessed by an expert via the DOPS method based on a 1 to 10 Likert scale for each mini skill. All the surgeries were also recorded and later evaluated via the VOPS method by an expert based on the same grading scale. Finally these grades were compared to each other, and the correlation between VOPS and DOPS grades was analyzed.

In order to minimize the inter-observer error, all VOPS and DOPS assessments were done by a single ophthalmology professor. To evaluate the intra-observer reproducibility of the VOPS assessment, all the videos were re-evaluated by the same observers within two weeks.

The absolute agreement of the grades was analyzed by one-way mixed effect model. The absolute agreement of a single observer's measurements was calculated with the intraclass correlation coefficient (ICC) from a 2-way mixed effect model. Comparisons between two groups were performed using Mann–Whitney U test. Measurements between the two methods were also compared using Bland–Altman analysis, which calculates the mean and 95% limit of agreement. Data were analyzed using SPSS software (version 18 for Windows; SPSS Inc., Chicago, IL, USA).

Results

There were 3 female and 7 male residents in our study group. Residents have been done a mean of 14.9 ± 3.5 (range 10–20) independent trabeculectomy before the assessments. DOPS grade was positively associated with number of independent trabeculectomy during glaucoma rotation (β=0.227, $p = 0.004$) However, DOPS grade was not correlated with the gender of residents ($\beta = -0.143$, $p = 0.85$).The intra-observer reproducibility of VOPS grades was 0.847 (95% CI: 0.634, 0.961).

We observed a linear relationship between VOPS and DOPS scores (Fig. 1). VOPS scores were consistently lower than DOPS scores (8.4 ± 1.08 vs. 8.9 ± 0.99, $p = 0.02$), as demonstrated by most points falling below the line of equality. The regression coefficient ($\beta = 0.970$, 95% CI: 0.57, 1.37) was not statistically different from the line of equality (slope of 1). There was a bias in which VOPS underestimates the score by 0.20 points compared to DOPS.

In Bland–Altman analysis, all data points fell within the 95% limits of agreement (−1.46, 0.46) (Fig. 2). The mean difference between VOPS and DOPS methods was −0.5 points across the range of mean scores (7.3–10 points), without variation of score difference at different mean score values. Fig. 2 illustrates the mean difference, 95% limits of

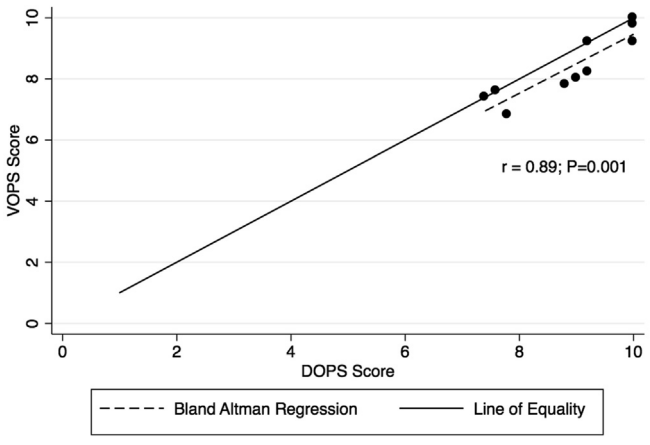

Fig. 1. Scatter plot of VOPS (y-axis) vs. DOPS (x-axis). There is a "line of equality" with slope of 1, representing perfect equivalence of the methods. Our points fall below the line of equality, suggesting that VOPS has negative bias. The linear regression for the points has a slope that is not significantly different from 1 (good equivalence). The y-intercept (constant term) of −0.20 shows that VOPS consistently scores residents 0.20 points lower vs. DOPS.

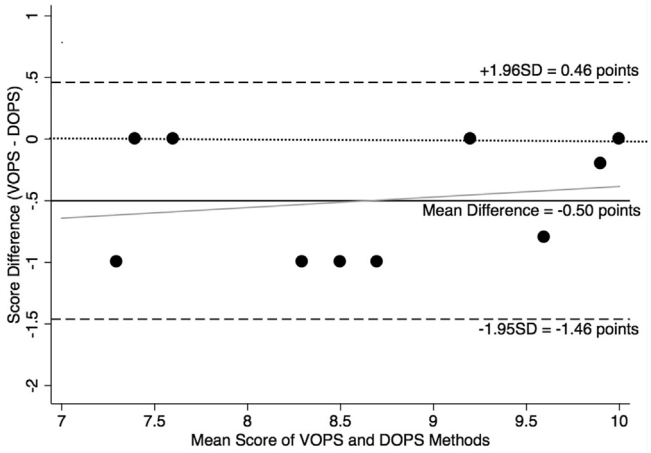

Fig. 2. Conventional Bland–Altman plot. The mean difference and 95% limits of agreement are shown.

agreement, and the linear regression describing the method agreement between VOPS and DOPS.

The result of univariate regression shows that VOPS grade was also associated with number of previous trabeculectomy (β=0.261, $p = 0.001$), but not the gender (β=0.001, $p > 0.99$).

Discussion

The traditional mentorship model of teaching and learning in medicine involving apprenticeship with a more experienced and senior expert is being challenged.[9] Recently, medical education is increasingly dynamic as technology shapes and guides medical educational policies, curricula, and assessment methods.[1,9] With these technological advancements, medical are developing new curricula, providing advanced educational tools, and introducing new methods of teaching and assessment.[2,3] Procedural and surgical skills are among the most important skills that medical students and residents should learn. Valid, reliable and efficient assessment methods for technical and surgical abilities are necessary given the growing number of medical students and residents.[4]

Direct observation of procedural skills (DOPS) is the most commonly used assessment method for evaluating these abilities.[7] Despite high face validity and closely resembling real clinical settings, there are problems with this method that can limit its effectiveness in assessment of surgical skills in newer educational systems.[9] The most important challenge is that multiple trainees cannot be accurately assessed at the same time by the limited number of clinical professors; with growing numbers of residents, this method is becoming less feasible. Another important shortcoming to the DOPS approach is that observers cannot be masked to the identity of the trainee being graded, introducing risk of observer bias. Additionally, while being acutely aware of being observed, trainees' performance may not accurately represent their usual performance during the evaluation. Newer, more systematic methods of assessment of surgical skills in teaching hospitals are needed to combat these challenges.[9]

VOPS is one recently introduced method of assessment. This new approach has the potential to be more objective, precise, and flexible. Any new assessment method requires demonstration of its validity, reliability, and feasibility and comparison with current assessment methods.[10] To our knowledge, we are the first group to evaluate VOPS reliability, validity, and feasibility in comparison to DOPS in ophthalmology surgical skills. Limited studies eist investigating this issue in other procedural specialties, including radiology and colonoscopy.[7,9]

In this study, we aimed to assess the agreement between grades of ophthalmology residents performing trabeculectomy measured through VOPS and DOPS assessment methods.

The reliability of the VOPS method was evaluated by grading the videos twice during a 2 week interval. Intra-observer reproducibility of VOPS grades was 0.847 (95% CI: 0.634–0.961), indicating acceptable reliability with this method.

The mean VOPS grades was significantly lower than the mean DOPS grades (8.4 vs. 8.9, $p = 0.02$). One explanation may be that the VOPS method provides the assessor with adequate time and a proper setting for precise, step-by-step assessment of the procedure, allowing more meticulous detection of any errors. However, a strong correlation was found between the VOPS and DOPS grades ($r = 0.89$, $p = 0.001$), suggesting that the two approaches are generally comparable. We therefore propose that the VOPS method can be used as an alternative to DOPS when assessing ophthalmologic surgical skills. VOPS is more systematic and can be blinded. It also allows the residents to get involved in the assessment process and compare their procedural skills with other trainees. An active role in the evaluation process may help trainees be more reflective of their performance and use feedback from professors and peers for self-improvement.

Several studies have demonstrated the feasibility, reliability, and validity of VOPS in surgical training at the resident level. In a study by Aggarwal et al, 19 novice and experienced surgeons performed laparoscopic cholecystectomies, and their intraoperative videos were graded using different scoring systems.[10] VOPS assessment using a global rating scale was able to distinguish between the novice and experienced surgeons. The VOPS grades showed adequate inter-rater reliability.

Driscoll et al also compared VOPS to DOPS in evaluating 9 residents' tissue-handling skills, using three scoring systems at two time points in surgical training.[12] Both VOPS and DOPS were reliable, demonstrating ICCs over 0.69. VOPS had adequate construct validity by detecting differences between trainees and experienced surgeons, including distinguishing trainees at different level of training, while DOPS was unable to detect the latter differences. There was also good interrater reproducibility of scores as well as strong correlation between test-retest scores across multiple scoring systems, supporting concurrent validity.

In a study by Dath et al, general surgeons graded 29 surgical residents performing laparoscopic operations on anesthetized pigs.[13] Mean interrater reliability of VOPS scores was

0.74. Reproducibility of scores was similar for both a global rating scale and a procedure-specific checklist. The scores correlated with the resident's level of training, though these differences did not reach statistical significance likely due to small sample size. Study investigators also reported an 80% reduction in assessment time compared to direct observation, since graders could fast-forward the video footage at viewer discretion. These findings support that VOPS is an attractive alternative to DOPS in resident evaluation.

However, there remain some challenges to the adoption of VOPS as an alternative to direct observation. In a study evaluating 22 surgical residents performing laparoscopic cholecystectomy, Scott et al compared these two approaches while using a global rating scale.[14] The correlation between VOPS and DOPS scores for five criteria was poor (<0.33), and DOPS scores were more reliable than VOPS scores (ICC 0.57 vs. 0.28). DOPS identified a score difference between residents with additional training compared to a control group, while VOPS did not detect any difference. Beckmann et al also investigated the feasibility of VOPS in obstetrics-gynecology residency training.[15] Of 50 procedures, only 23 videos could be generated while the rest suffered from camera problems. Though it was feasible to generate scores based on checklist items and there was the benefit of 34% reduction in assessment time, score agreement between graders varied widely. Overall, these suggest that though VOPS has many benefits, inadequate technological support in hospitals may be a barrier to its utilization, though many advancements have been made since the latter two studies were conducted. Additionally, further work is required to assess the validity of video-based assessments when multiple graders are involved, and its utility may vary across surgical subspecialties given their different procedural skills. The use of standardized checklists may be helpful.

This study has several limitations. We had a small sample size given the pilot nature of this study, and future investigations with larger study populations may be considered to validate our findings. Additionally, we were unable to assess inter-grader reproducibility of VOPS grades as only one grader was utilized. Finally, we did not assess residents' and assessors' satisfaction regarding the VOPS approach.

Though traditionally VOPS has had lower face validity than DOPS, its many advantages make it an acceptable and even desirable alternative for assessment of procedural skills in medical training. Our study presents data demonstrating the feasibility, reliability, and validity of VOPS compared to DOPS in the evaluation of ophthalmology residents' surgical skills. In summary, the results of the present study showed that VOPS is a feasible, valid and reliable assessment method for procedural skills of medical students and residents that can be used as an alternative for DOPS method. However, VOPS might underestimate DOPS in evaluating surgical skills of residents.

References

1. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003−2009. *Med Educ*. 2010;44:50−63.
2. Vozenilek J, Huff JS, Reznek M, Gordon JA. See one, do one, teach one: advanced technology in medical education. *Acad Emerg Med*. 2004;11:1149−1154.
3. Cox M, Irby DM, Reznick RK, MacRae H. Teaching surgical skills—-changes in the wind. *N Engl J Med*. 2006;355:2664−2669.
4. Hauer KE, Hodgson CS, Kerr KM, Teherani A, Irby DM. A national study of medical student clinical skills assessment. *Acad Med*. 2005;80:S25−S29.
5. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002;287:226−235.
6. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med*. 1990;65:S63−S67.
7. Barton JR, Corbett S, van der Vleuten CP, Programme EBCS. The validity and reliability of a direct observation of procedural skills assessment tool: assessing colonoscopic skills of senior endoscopists. *Gastrointest Endosc*. 2012;75:591−597.
8. Shah GS, Pouladi A, Bahram RM, Farhadifar F, Khatibi R. Evaluation of the effects of direct observation of procedural skills (DOPS) on clinical externship students' learning level in obstetrics ward of Kurdistan University of Medical Sciences. *J Med Educ Winter Spring*. 2009;13:29−33.
9. Bari V. Direct observation of procedural skills in radiology. *Am J Roentgenol*. 2010;195:W14−W18.
10. 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;247:372−379.
11. Schacknow PN, Samples JR. *The Glaucoma Book: A Practical, Evidence-based Approach to Patient Care*. New York: Springer; 2010.
12. Driscoll PJ, Paisley AM, Paterson-Brown S. Video assessment of basic surgical trainees' operative skills. *Am J Surg*. 2008;196:265−272.
13. Dath D, Regehr G, Birch D, et al. Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Surg Endosc*. 2004;18:1800−1804.
14. Scott DJ, Rege RV, Bergen PC, et al. Measuring operative performance after laparoscopic skills training: edited videotape versus direct observation. *J Laparoendosc Adv Surg Tech*. 2000;10:183−190.
15. Beckmann RB, Lipscomb GH, Ling FW, Beckmann CA, Johnson H, Barton L. Computer-assisted video evaluation of surgical skills. *Obstet Gynecol*. 1995;85:1039−1041.