

Direct Observation vs. Video-Based Assessment in Flexible Cystoscopy

Julia Dagnaes-Hansen, BMSc,^{*,†,‡} Oria Mahmood, MD,^{*,†} Sarah Bube, MD,^{*,§} Flemming Bjerrum, PhD,^{||} Yousif Subhi, MD,^{†,¶} Malene Rohrsted, MD,^{*,‡} and Lars Konge, PhD^{*,†}

^{*}Copenhagen Academy for Medical Education and Simulation, Copenhagen, Capital Region, Denmark; [†]Faculty of Health and Medical Sciences, Faculty of Medicine, University of Copenhagen, Capital Region, Denmark; [‡]Department of Urology, University Hospital of Copenhagen, Rigshospitalet, Copenhagen, Denmark; [§]Department of Urology, Zealand University Hospital, Roskilde, Denmark; ^{||}Department of Surgery, Herlev Gentofte Hospital, University Hospital of Copenhagen, Herlev, Denmark; and [¶]Department of Ophthalmology, Zealand University Hospital, Roskilde, Denmark

OBJECTIVE: Direct observation in assessment of clinical skills is prone to bias, demands the observer to be present at a certain location at a specific time, and is time-consuming. Video-based assessment could remove the risk of bias, increase flexibility, and reduce the time spent on assessment. This study investigated if video-based assessment was a reliable tool for cystoscopy and if direct observers were prone to bias compared with video-raters.

DESIGN: This study was a blinded observational trial. Twenty medical students and 9 urologists were recorded during 2 cystoscopies and rated by a direct observer and subsequently by 2 blinded video-raters on a global rating scale (GRS) for cystoscopy. Both intrarater and interrater reliability were explored. Furthermore, direct observer bias was explored by a paired samples *t*-test.

RESULTS: Intrarater reliability calculated by Pearson's *r* was 0.86. Interrater reliability was 0.74 for single measure and 0.85 for average measures. A hawk-dove effect was seen between the 2 raters. Direct observer bias was detected when comparing direct observer scores to the assessment by an independent video-rater (*p* < 0.001).

CONCLUSION: This study found that video-based assessment was a reliable tool for cystoscopy with 2 video-raters. There was a significant bias when comparing direct observation with blinded video-based assessment. (J Surg Ed 1:111-114. © 2017 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: cystoscopy, video recording, interrater variability, rater-based assessment, surgical education

COMPETENCIES: Patient Care, Practice-Based Learning and Improvement

INTRODUCTION

With the growing use of competency-based training in medical education, it is important to determine optimal strategies for performance assessment.¹⁻⁵ There are multiple ways of assessing trainees in the daily clinical setting with direct observation as the standard method. However, time spent on the traditional Halstedian apprenticeship is often limited, and assessing procedures can be prone to bias,⁶ challenging, and very time-demanding, especially if the assessor has to be present during the procedure.

Advances in technology open the possibility of using video-based assessments instead of direct observation.⁷ In direct assessment, the rater can be biased by knowing the identity of the trainee^{8,9} due to several cognitive and social mechanisms such as social relations between the trainee and rater and previous experiences.^{5,10} Multiple biases can influence the overall assessment in direct observation. Examples of such are anchoring bias where the assessor holds on to an initial opinion or observation of the trainee and visceral bias where the assessment is based on emotions rather than objective data.¹¹ Furthermore, a halo-effect is often seen, where the subsequent procedures are affected if the assessor has had a good or bad first-hand impression.¹² The halo-effect is seen in both direct assessment as well as video-based assessment but could possibly be diminished in the latter by letting the participant perform 2 or more consecutive procedures.

Another benefit of video-based assessment is that several observers can rate the same video in order to improve

Correspondence: Inquiries to Julia Karina Dagnaes-Hansen, BMSc, Copenhagen Academy for Medical Education and Simulation, CAMES, Frederik V's vej 11, 2100, Copenhagen, Denmark; e-mail: julia.dagnaes@gmail.com

reliability.¹³ In addition, video-based assessments are proven less time-consuming than direct observation. Studies examining video-based assessment of laparoscopy have shown raters to spend 34% to 80% less time on the video-rating process compared with direct observation.^{14,15}

Video-rating has previously been used in other procedures and settings such as laparoscopic surgery with good results.¹⁵⁻¹⁸ A limited number of studies have investigated video-based assessment in simpler and shorter procedures, where the overall results were successful.^{8,19,20} Cystoscopy is one of the first procedures to be learned in the urologic residency curriculum and is therefore a relevant object of investigation for video-rating.

The objective of this study was to investigate whether video-based rating is a reliable tool for assessing cystoscopy performance. Furthermore, we sought to investigate if direct observers are biased (such as visceral or anchoring bias¹¹) compared to video-based assessors.

METHODS

Setting and Participants

Medical students from the Faculty of Health and Medical sciences at the University of Copenhagen and experienced urologists, who had performed more than 100 cystoscopies within the last year, performed flexible cystoscopies at the outpatient clinic of the Department of Urology, Copenhagen University Hospital Rigshospitalet, Denmark. All cystoscopies were performed on patients already scheduled for cystoscopy and were performed under local anesthesia. All participants gave written consent prior to participating. The study submitted to the Danish Ethical Research Committee (H-15011265) which assessed the study and found that no approval was necessary.

Patients with anatomical abnormalities or previous reconstructive bladder surgery were excluded. All participants performed 2 consecutive procedures.

Direct Assessment and Recording

An experienced urologist introduced the scope into the urethra in order to prevent unnecessary inconvenience for the patients. When in the bladder, the novices took over the handling of the scope. The urologist participants introduced the scope in the bladder and therefore the recording was not started until the scope was in the bladder, as this part of the procedure was not rated.

The performance was scored by a direct observer using the Global rating score (GRS) (Appendix) for cystoscopy.²¹ The direct observer was the same assessor throughout the entire study. Points from 1 to 5 on a 5-point Likert scale were given on the 5 parameters of the GRS. The identity of the participant was known to the direct observer. The video-recordings started with the cystoscope entering the

bladder and stopped at the end of the procedure when the cystoscope was removed through the urethra. The procedure was recorded directly through the lens of the cystoscope, hence the identity of the performer was therefore completely concealed on the video.

Raters and Rater Training

The 2 raters were both experienced urologists who had performed more than 100 cystoscopies. To minimize construct-irrelevant variance, we used rater training in order to increase accuracy of assessments.^{1,22} After the direct observation and recording of the procedures, the 2 raters met and individually watched and assessed 2 videos. Afterwards they compared results and discussed their ratings in order to ensure agreement on the scoring system. The videos were randomized by a third person, who had not been present at the recordings, to a new number using <http://www.random.org/> and put into a secured web-based video-rating software²³ so the direct observer would be blinded in the following video assessment. Thereafter, the direct observer (DO) assessed the videos again blinded as a video-based assessor (VBA1) to be able to assess both interrater as well as intrarater reliability. Video-based assessment was delayed a month from the last video being recorded to minimize the risk of the direct observer remembering the procedures. Another experienced urologist (VBA2) assessed the video-recorded performances.

Data Collection and Analysis

Intrarater reliability was explored by comparing the direct observer's assessments under direct observation with the same rater's assessments based on video-recordings (DO vs. VBA1). Interrater reliability was explored by comparing assessments based on video-recordings (VBA1 vs. VBA2). Reliability measures were calculated using Intraclass Correlation Coefficients (ICC), single and averages measures.

To establish whether there was a Hawk-Dove effect between the 2 assessors, an effect seen where some assessors are more stringent and other more lenient, we calculated a paired samples *t*-test comparing ratings given by the 2 different video-raters.

Anchoring bias was explored by calculating the difference between direct observation and video-based observation and comparing these delta-values of novices and experienced, respectively. An independent samples *t*-test was used to make this comparison. Differences were considered statistically significant when the *p* value was <0.05.

A pass/fail standard was set by using the contrasting groups' standard setting method, and the numbers of false-positive and false-negative results were explored.

We conducted a reliability analysis and compared the Cronbach α for direct assessment and for video-based

Download English Version:

<https://daneshyari.com/en/article/8834699>

Download Persian Version:

<https://daneshyari.com/article/8834699>

[Daneshyari.com](https://daneshyari.com)