ORIGINAL ARTICLE

Assessment of competence for caesarean section with global rating scale

Rahat Najam Qureshi, Syeda Kauser Ali

Abstract

Objective: To establish as reliable and valid the nine-point global rating scale for assessing residents' independent performance of Caesarean Section.

Methods: The validation study was conducted at the Department of Obstetrics and Gynaecology, Aga Khan University Hospital, from April to December 2008, and comprised 15 residents during 40 Caesarean Sections over 9 months. Independently two evaluators rated each procedure and the difficulty of each case.

Results: The observations per faculty ranged from 1-8 (mean 4.07± 2.56). The Year 4 residents were observed the most i.e. 32 (40%), followed by Year 3, 30 (37.5%); Year 2; 14 (17.5%); and Year 1, 4 (5%). Mean time required for observation of the surgery was 43.81±14.28 (range: 20-90) with a mode of 45 min. Mean aggregate rating on all items showed gradual progression with the year of residency. The assessment tool had an internal consistency reliability (Cronbach's alpha) of 0.9097 with low inter-rater reliability.

Conclusion: The evaluation tool was found to be reliable and valid for evaluating a resident's competence for performing Caesarean Section. Training of the assessors is required for a better inter-rater agreement.

Keywords: Assessment, Competence, Caesarean section, Rating scale. (JPMA 63: 1003; 2013)

Introduction

Assessment plays a vital role in ensuring the development of appropriate skills and professional attributes. In addition, it also offers information to students about the knowledge, skills and other attributes they can expect to possess after successfully completing an academic programme. It also establishes ways for teachers and assessors to understand the dimensions of student learning when seeking to improve student achievement and the educational process. Teachers and assessors can then alter the teaching and/or assessment style to become more focussed and achieve higher level of efficiency in achieving the educational objectives.

Disciplines which require trainees to learn surgical skills have historically lacked objective assessment.[1] Trainees of these programmes are required to maintain a log of the educational opportunities and experiences of surgical procedures. This is then used to evaluate the quality, variety and number of operative procedures completed either as an assistant or independently. Since the introduction of Objective Structured Clinical Examination (OSCE), psychomotor skill assessment has become a routine part of Clinical Skills Assessment. Yet assessment of operative skills has lagged behind.[2,3] Surgeons have explored various options which could come close to the real thing. These include Bench station assessment which tests technical skills by observing residents while they perform on animal tissues and special models.[4] Innovative computer technology[5] has revolutionised teaching/learning and assessment, but it cannot be easily accessed by all special programmes which operate in resource-constrained environment. Direct Observation of Procedural Skills[6] (DOPS) which is an effort to formalise the teaching and assessment of surgical procedures has been found to be useful, but has not been applied on a large scale as yet.

Caesarean Section (CS) which is a common procedure in obstetrics and gynaecology comprises complex skills backed with knowledge and embellished with professional attitudes. Residents initially observe and then assist, gradually becoming competent to perform the CS procedure independently. It usually takes first two years of a residency programme for the residents to perform the complete procedure independently, but under supervision.

The supervisors use checklist or rating scales to assess the readiness of the resident to perform the procedure independently. Studies on assessment have provided evidence that checklists are not very reliable as they promote thoroughness over efficiency.[7] In addition one of the major concerns of assessment is not on how many steps were done but if they were done well. It is appreciated that as the learner progresses through the years of residency education from being a novice to competent, he/she takes shortcuts based on prior learning and experiences. These shortcuts are accepted as a move towards expertise.[8]

Global rating scale (GRS) is considered to be valid and reliable

Department of Obstetrics and Gynaecology, Aga Khan University Hospital, Karachi, Pakistan.

Correspondence: Rahat Najam Qureshi. Email: rahat.qureshi@aku.edu

form of assessment for operating room performances of residents.[1,4,7] The steps of the specific skill are listed as items in a column against which marks are given on a (5-7-9-point) rating scale, with the middle and the extreme points anchored by explicit descriptors to help in the criterion referenced assessment of performance.[1,4,7,9,10] A universal GRS for the evaluation of technical skills in the Operating Room (OR) bears good correlation with the residents' ability in clinical areas, including surgical ability on the bench.[1] When scored by specialists, the results have shown to have a higher inter-station reliability, better construct validity, and better concurrent validity in comparison with checklists.[1,10]

The current study was done to determine the reliability and validityof GRS to assess the competence of obstetrics and gynaecology residents for CS.

Subjects and Methods

The validation study was conducted at the Department of Obstetrics and Gynaecology (Obgyn), Aga Khan University (AKU) Hospital, Karachi, from April to December 2008. The study was approved by the ethics review committee and the faculty and residents were briefed and their consent was taken beforehand.

The department had a total of 18 residents in 2008. As part of the inclusion criteria we only allowed those residents to perform C-section with whom the faculty felt satisfied that they had the capacity to perform the procedure safely while under direct supervision. Fifteen (83.33%) residents in the Obgyn residency programme from Years 1 to 4 were eligible for the study as long as the primary consultant felt them to be competent enough to be allowed to operate. The observations were only done during daytime elective CS. The attending consultant participated voluntarily and as per the policy of the department either scrubbed with the resident or was available in the OR. All the patients included in the study had signed the in-patient informed consent form which allowed the attending consultant/faculty or appointed delegate to operate.

Each resident was observed by two consultants during the CS. Each evaluator used the GRS form independently to rate the resident's performance during the procedure. The assessors included both full-time and part-time faculty members.

The assessors were asked not to interrupt during the procedure, but were allowed to informally provide verbal feedback after observation and submission of the completed GRS form. The forms were not used as part of the formal evaluation system.

A total of 80 observations were made for the 15 residents who performed CS on 40 patients. Information regarding the number of times each resident had assisted CS, performed CS under supervision or performed CS independently was obtained from the logbooks.

The GRS form had 10 items for the steps relevant to the procedure and one item for overall ability to be observed on a seven-point scale for scoring the resident's performance on each item. The items listed in the form have been identified as critical elements in the performance of the surgical procedure[1,6] and were agreed upon by the supervisors, faculty and the consultants. The items included: Appropriate pre-operative preparations; Follows aseptic technique; Identifies relevant structures; Follows sequence in technique; Handles tissue gently; Uses instruments effectively; Uses appropriate sutures; Delivers baby with care; Deals with unexpected events and seeks help where needed; Clears post-operative instructions; and Overall ability to perform the procedure.

Item number 8 relating to delivery of the baby was introduced upon the recommendation of all the faculty members from Obgyn and General Surgery at AKUH. The raters were also asked to identify the difficulty level for each procedure for the resident's with respect to the year of training.

Results

A total of 15 residents were evaluated by 16 evaluators over 40 CS operations. The number of GRS-CS forms filled per resident ranged from 2 to 8 (mean 7.45±3.92; mode 8.00). The number of observations per faculty ranged from 1 to 8 with a mean of 4.07±2.56.

The number of forms filled for each year of residency training showed that the residents in postgraduate year (PGY) 4 were observed the most times followed by PGY3 (Table-1).

Time taken to complete the observation was 20-90 minutes, with a mean time of 43.81±14.28; mode 45 min i.e. 38 (47%) observations took 40-45 minutes. Fifteen (93.75%) faculty evaluators were satisfied with the form while 1 (6.25%) was highly dissatisfied. Faculty also rated each CS on its level of difficulty on a three-point scale i.e. easy, moderately difficult, and high level of difficulty. This was a subjective assessment based on the faculty's judgment of the case in regard to the

Table-1: Residents' distribution with respect to year of residency.

Year of residency	Number of residents	Number of forms filled (%)
4	5	32(40)
3	5	30(37.5)
2	4	14(17.5)
1	1	4(5)
Total	15	80 (100%)

Table-2: Item analysis.

Items (n=80 forms)	Mean rating
Pre-operative preparation	3.69±3.01
Follows aseptic technique	4.87±2.62
Identifies relevant structure	5.40±2.02
Follows sequence in technique	5.30±2.25
Handles tissue gently	5.41±2.03
Effectively uses instruments	5.12±2.11
Uses appropriate sutures	5.45±2.09
Delivers baby with care	4.47±2.78
Deals with any unexpected event	4.13±2.73
Clear post-procedure instruction	5.23±2.58

year of residency training the candidate was in. The mean difficulty of the cases was 1.4±0.544, showing that the cases ranged from being easy to moderately difficult for the year of residency training. This confirms the general perspective that faculty will not give highly difficult cases to the residents as the first operator.

The mean ratings on each item ranged from 3.69 to 5.45 with 'pre-operative preparation' getting the least and 'using appropriate sutures' being awarded the highest mean score (Table-2).

All items correlated positively with each other except for the item 'deals with any unexpected event during the procedure' which correlated negatively with aseptic techniques (Table-3).

The instrument (GRS form) tool had an internal consistency reliability (Cronbach's alpha) of 0.9097 with an inter-rater reliability of 0.176. Content validity was ensured by obtaining agreement from all the teachers, assessors and supervisors of the residents regarding the items to be used on the GRS form. Evidence of construct validity was obtained by performing factor analysis of the GRS form using varimax rotation. Four factors were identified on Principal component analysis - pre-operative preparation, operative technique, post-operative skills and safe delivery (Table-4).

Further evidence of construct validity was determined by examining the discriminatory power of the instrument between different levels of residency training. The mean rating given by evaluators as per the year of Residency on all items showed gradual progression with the year of residency (Figure-1). Independent Kruskal Wallis test showed significant ($p<0.01$) progression of ranking over the years. Mann Whitney test showed that the difference was not significant between years two and three.

The most dramatic difference was seen in the item titled 'delivers baby with care' in which Year 1 and Year 4 residents had a significant different score ($p<0.01$).

The ratings given by the assessors showed that the senior

Table-3: Globar Rating Scale - inter item correlation (p values in parenthesis).

GRS Items	Pre op preparation	Aseptic technique	Identifies relevant structure	Sequence of technique	Handles tissue gently	Effectively uses instruments	Uses appropriate sutures	Delivers baby with care	Deals with any unexpected event	Clear post procedure instructions	Overall procedural skills
Pre op preparation	1.000										
Aseptic technique	0.488 (0.000)	1.000									
Identifies relevant structures	0.085 (0.225)	0.24 (0.016)	1.000								
Sequence of technique	0.24 (0.016)	0.3 (0.003)	0.721 (≤0.0001)	1.000							
Handles tissue gently	0.181 (0.054)	0.185 (0.051)	0.703 (≤0.0001)	0.717 (≤0.0001)	1.000						
Effectively uses instruments	0.204 (0.035)	0.179 (0.056)	0.625 (≤0.0001)	0.657 (≤0.0001)	0.695 (≤0.0001)	1.000					
Uses appropriate sutures	0.103 (0.181)	0.161 (0.076)	0.657 (≤0.0001)	0.652 (≤0.0001)	0.698 (≤0.0001)	0.761 (≤0.0001)	1.000				
Delivers baby with care	0.14 (0.108)	0.156 (0.084)	0.347 (≤0.0001)	0.425 (≤0.0001)	0.466 (≤0.0001)	0.398 (≤0.0001)	0.488 (≤0.0001)	1.000			
Deals with any unexpected event	0.064 (0.287)	-0.100 (0.188)	0.026 (0.41)	0.123 (0.139)	0.173 (0.062)	0.131 (0.123)	0.08 (0.241)	0.17 (0.066)	1.000		
Clear post procedure	0.09 (0.213)	0.047 (0.34)	0.185 (0.05)	0.259 (0.01)	0.265 (0.009)	0.332 (0.001)	0.337 (0.001)	0.305 (0.003)	0.253 (0.012)	1.000	
Candidate procedural skills	0.178 (0.058)	0.176 (0.059)	0.738 (≤0.0001)	0.73 (≤0.0001)	0.814 (≤0.0001)	0.718 (≤0.0001)	0.758 (≤0.0001)	0.641 (≤0.0001)	0.123 (0.138)	0.344 (0.001)	1.000

Table-4: Factor analysis.

Items	Factors			
	Pre-operative preparation	Operative technique	Post-operative skills	Safe delivery
Pre-operative preparation	0.962			
Aseptic technique	0.52			
Identifies relevant structures		0.832		
Follows sequence in technique		0.827		
Hand tissue gently		0.772		
Effectively use instruments		0.744		
Uses appropriate sutures		0.792		
Delivers baby with care				0.872
Deals with unexpected event			0.855	
Clear post-procedure instructions			0.725	

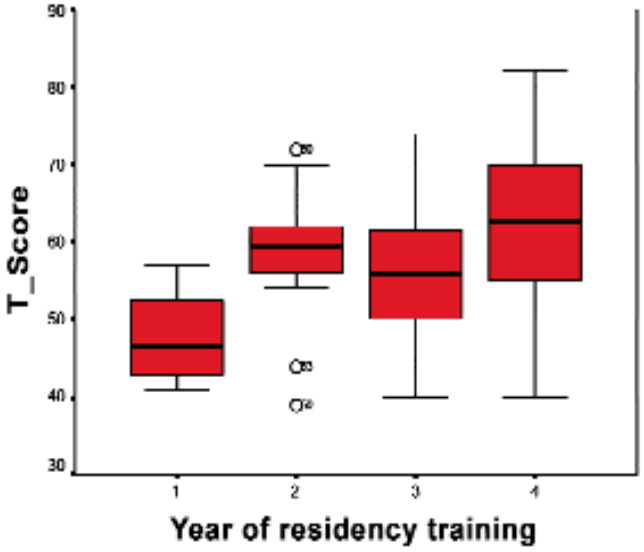

Figure-1: Mean total ratings according to year of residency education.

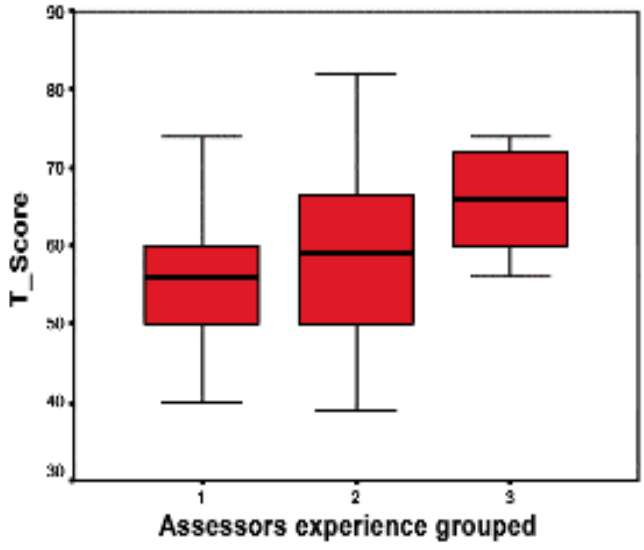

Figure-2: Mean rating by assessors grouped on the basis of years of experience.

faculty rated the residents higher than their junior counterparts (Figure-2). Since the number of part-time faculty was only 1, hence she was not included in the calculations.

Discussion

There is a move towards Work-Based Assessment for bringing in a more authentic evaluation. However, because of the complexity of the process involved in the assessment of surgical skills, this has proved to be difficult. Using models and computers, such methods have their utility, but they cannot replace real-life experiences and learning opportunities provided when operating on patients.

The Global Rating Scale is applied as an assessment tool in the workplace for the evaluation of surgical competence in CS. Despite different strategies for evaluating surgical procedures, their feasibility and reliability still remain unresolved.[3] The challenge is to develop assessment tools which can be standardised to take into account the experience of the assessor, the level of the trainee, the difficulty of the case, the complex interplay of cognitive/individual skills and other factors which may be influencing the observation and the performance.

Validity evidence confirms that the instrument differentiates between residency levels. Validation has yielded good results with discrimination between different levels.[10] The item, 'delivers baby with care' in this study behaved as a unique and essential item for assessment and encompassed many strategies and steps taken during the procedure. The wording of this item reveals the underlying principles and practices which are consciously or sub-consciously recognised and evaluated by the supervisor. Interestingly it correlates positively with most items.[1] There appears to be a difference in the individual skills which jointly facilitates in bringing satisfactory results in surgical competency. Some of the components may be relatively easier to master and may be more heavily based on cognitive skills of the individual

while some may be more heavily dependent on repeated attempts and practice.[11]

The internal structure of the test format demonstrated good evidence for the test items. There was good homogeneity in the items and a positive correlation. Principal factor analysis further positively defined the relationship of the different items with each other. The authenticity of the validity judgment in the form of internal structure was low as the observers were not trained to use the GRS. In order to keep the assessment as real as possible, the investigators on purpose ignored the training. Another reason for the low inter-rater reliability may be due to the differences in the number of years of teaching experience. Therefore the junior faculty members may have had a different marking pattern from the seniors.

Factor analysis showed that the instrument was assessing four distinct abilities, which could be identified as pre-operative preparation, surgical expertise, post-operative skills and safe delivery. It is important that supervisors were aware of the specific academic needs of the residents and taught and assessed them accordingly.

The study was confined to one hospital which may be considered as strength as it controls for extraneous factors which may influence the resident's performance evaluation and the evaluator. On the other hand, this restricted the number of residents who were observed and thus barred a more detailed analysis.

The study attempted to define a standardised process for certifying competence in CS. At present we need to know how many observations, total number of cases performed and diversity of case complexity to certify competence in surgical procedures in general and CS in particular.

This instrument needs to be investigated further to evaluate its role in promoting learning of the residents, and the development of their self -confidence. It may also assist in defining the critical items which indicate expertise in residents' performance. It is noted that no emergency cases were included and it is possible that surgical practice in an emergency situation involves other factors which should also be included in the assessment.

Considering the urgency for implementing programmes and strategies to meet the Millennium Development Goals (MDGs) some countries have trained medical assistants for working in rural settings. It is even more important in such setting to ensure that the trainees have acquired reasonable level of skills for safe practice when they work unsupervised or with little or no assistance from the senior attendant.

Conclusion

The validation study yielded good results with good discrimination between performances of residents at different levels of training. The GRS can be used for continuous assessment and recording the progress of a resident for different operative procedures. The use of the GRS for learning could be maximised by making constructive feedback a compulsory part of the process.

Acknowledgement

We are grateful to the Obstetrics and Gynaecology Departmental Research Committee, Aga Khan University, for funding the study, the faculty, the staff and the residents who participated in the study. We would also like to thank Ms. Humera Ismail for data management and Ms. Dania Anwar for editorial assistance.

References

1. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, et al. An objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg 1997; 84: 273-8.
2. Vaz F, Bergstrom S, da Luz Vaz Maria, Langa J, Bugalho A. Training medical assistants for surgery. Bull World Health Organ 1999; 77: 688-91.
3. Crofts TJ, Griffiths JM, Sharma S, Wygrala J, Aitken RJ. Surgical training: an objective assessment of recent changes for a single health board BMJ. 1997; 314: 891-5.
4. Reznick RK, Regehr G, Mac Rae H, Martin J, McCulloch W. Testing technical skills via an innovative "bench station" examination. Am J Surg 1997; 173: 226-30.
5. Sourin A, Sourina O, Howe TS. Virtual orthopedic surgery training. IEE Computer Graphics and Applications 2000; 20: 6-9.
6. Department of Health. DOPS (Direct Observation of Procedural Skills). In: Modernizing Medical Careers. London: DoH; 2005. (Online) (Cited 2007 June 8). Available from URL: http://www.mmc.nhs.uk/pages/assessment/dops.
7. 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.
8. Norcini JJ, Holmboe ES, Hawkins RE. Evaluation challenges in the era of outcomes-based education. In: Holmboe ES, Hawkins RE (eds). Practical Guide to the Evaluation of Clinical Competence. Philadelphia, PA: Mosby; 2008; pp 5-6.
9. MacRae HM, Vu NV, Graham B, Word-Sims M, Colliver J, Robbs R. Comparing checklists and databases with physicians' ratings as measures of students' history and physical-examination skills. Acad Med 1995; 70: 313-7.
10. Doyle JD, Webber EM, Sidhu RS. Universal global rating scale for the evaluation of technical skills in the operating room. Am J Surg 2007; 193: 551-5.
11. Reznick RK, MacRae H. Teaching surgical skills - changes in the wind. N Engl J Med 2006; 355: 2664-9.