Education and Training

Critical Analysis of Case Based Discussions

J M L Williamson and A J Osborne

Introduction

Assessment and evaluation are the foundations of learning; the former is concerned with how students perform and the latter, how successful the teaching was in reaching its objectives. Case based discussions (CBDs) are structured, non-judgmental reviews of decision-making and clinical reasoning[1]. They are mapped directly to the surgical curriculum and "assess what doctors actually do in practice"[1]. Patient involvement is thought to enhance the effectiveness of the assessment process, as it incorporates key adult learning principles: it is meaningful, relevant to work, allows active involvement and involves three domains of learning[2]:

- Clinical (knowledge, decisions, skills)
- Professionalism (ethics, teamwork)
- Communication (with patients, families and staff)

The ability of work based assessments to test performance is not well established. The purpose of this critical review is to assess if CBDs are effective as an assessment tool.

Validity of Assessment

Validity concerns the accuracy of an assessment, what this means in practical terms, and how to avoid drawing unwarranted conclusions or decisions from the results. Validity can be explored in five ways: face, content, concurrent, construct and criterion-related/predicative.

CBDs have high face validity as they focus on the role doctors perform and are, in essence, an evolution of 'bedside oral examinations'[3]. The key elements of this assessment are learnt in medical school; thus the purpose of a CBD is easy for both trainees and assessors to validate[1]. In terms of content validity, CBDs are unique in assessing a student's decision-making and which, is key to how doctors perform in practice. However, as only six CBDs are required a year, they are unlikely to be representative of the whole curriculum. Thus CBDs may have a limited content validity overall, especially if students focus on one type of condition for all assessments.

Determining the concurrent validity of CBDs is difficult as they assess the pinnacle of Miller's triangle – what a trainee 'does' in clinical practice (figure1)[4]. CBDs are unique in this aspect, but there may be some overlap with other work based assessments particularly in task specific skills and knowledge. Simulation may give some concurrent validity to the assessment of judgment. The professional aspect of assessment can be validated by a 360 degree appraisal, as this requests feedback about a doctor's professionalism from other healthcare professionals[1].
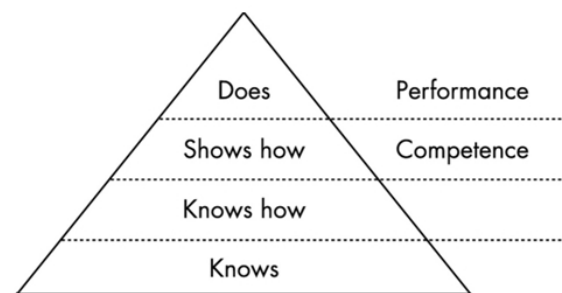

Figure 1: Miller's triangle[4]

CBDs have high construct validity, as the assessment is consistent with practice and appropriate for the working environment. The clinical skills being assessed will improve with expertise and thus there should be 'expert-novice' differences on marking[3]. However the standard of assessment (i.e. the 'pass mark') increases with expertise – as students are always being assessed against a mark of competency for their level. A novice can therefore score the same 'mark' as an expert despite a difference in ability.

In terms of predictive validity performance-based assessments are simulations and examinees do not behave in the same way as they would in real life[3]. Thus, CBDs are an assessment of competence ('shows how') but not of true clinical performance and one perhaps could deduct that they don't assess the attitude of the trainee which completes the cycle along with knowledge and skills ('does')[4]. CBDs permit inferences to be drawn concerning the skills of examinees that extend beyond the particular cases included in the assessment[3]. The quality of performance in one assessment can be a poor predictor of performance in another context. Both the limited number and lack of generalizability of these assessments have a negative influence on predictive validity[3].

Reliability of Assessment

Reliability can be defined as "the degree to which test scores are free from errors of measurement". Feldt and Brennan describe the 'essence' of reliability as the "quantification of the consistency and inconsistency in examinee performance"[5]. Moss states that less standardized forms of assessment, such as CBDs, present serious problems for reliability[6]. These types of assessment permit both students and assessors substantial latitude in interpreting and responding to situations, and are heavily reliant on assessor's ability. Reliability of CBDs is influenced by the quality of the rater's training, the uniformity of assessment, and the degree of standardization in examinee.

Rating scales are also known to hugely affect reliability – understanding of how to use these scales must be achieved by all trainee assessors in order to achieve marking consistency. In CBD assessments, trainees should be rated against a level of completion at the end of the current stage of training (i.e. core or higher training)[1]. While accurate ratings are critical to the success of any WBA, there may be latitude in the interpretation of these rating scales between different assessors. Assessors who have not received formal WBA training tend to score trainees more generously than trained assessors[7-8]. Improved assessor training in the use of CBDs and spreading assessments throughout the student's placement (i.e. a CBD every two months) may improve the reliability and effectiveness of the tool[1].

Practicality of Assessment

CBDs are a one-to-one assessment and are not efficient; they are labour intensive and only cover a limited amount of the curriculum per assessment. The time taken to complete CBDs has been thought to negatively impact on training opportunities[7]. Formalized assessment time could relieve the pressure of arranging ad hoc assessments and may improve the negative perceptions of students regarding CBDs.

The practical advantages of CBDs are that they allow assessments to occur within the workplace and they assess both judgment and professionalism – two subjects on the curriculum which are otherwise difficult to assess[1]. CBDs can be very successful in promoting autonomy and self-directed learning, which improves the efficiency of this teaching method[9]. Moreover, CBDs can be immensely successful in improving the abilities of trainees and can change clinical practice – a feature than is not repeated by other forms of assessment[8].

One method for ensuring the equality of assessments across all trainees is by providing clear information about what CBDs are, the format they take and the relevance they have to the curriculum. The information and guidance provided for the assessment should be clear, accurate and accessible to all trainees, assessors, and external assessors. This minimizes the potential for inconsistency of marking practice and perceived lack of fairness[7-10]. However, the lack of standardization of this assessment mechanism combined with the variation in training and interpretation of the rating scales between assessors may result in inequality.

Formative Assessment

Formative assessments modify and enhance both learning and understanding by the provision of feedback[11]. The primary function of the rating scale of a CBD is to inform the trainee and trainer about what needs to be learnt[1]. Marks per see provide no learning improvement; students gain the most learning value from assessment that is provided without marks or grades[12]. CBDs have feedback is built into the process and therefore it can given immediately and orally. Verbal feedback has a significantly greater effect on future performance than grades or marks as the assessor can check comprehension and encourage the student to act upon the advice given[1,11-12]. It should be specific and related to need; detailed feedback should only occur to help the student work through misconceptions or other weaknesses in performance[12]. Veloski, et al, suggests that systemic feedback delivered from a credible source can change clinical performance[8].

For trainees to be able to improve, they must have the capacity to monitor the quality of their own work during their learning by undertaking self-assessment[12]. Moreover, trainees must accept that their work can be improved and identify important aspects of their work that they wish to improve. Trainee's learning can be improved by providing high quality feedback and the three main elements are crucial to this process are[12]:

- Helping students recognise their desired goal
- Providing students with evidence about how well their work matches that goal
- Explaining how to close the gap between current performance and desired goal

The challenge for an effective CBD is to have an open relationship between student and assessor where the trainee is able to give an honest account of their abilities and identify any areas of weakness. This relationship currently does not exists in most CBDs, as studies by Veloski, et al[8] and Norcini and Burch[9] who revealed that only limited numbers of trainees anticipated changing their practice in response to feedback data. An unwillingness to engage in formal self-reflection by surgical trainees and reluctance to voice any weaknesses may impair their ability to develop and lead to resistance in the assessment process. Improved training of assessors and removing the scoring of the CBD form may allow more accurate and honest feedback to be given to improve the student's future performance. An alternative method to improve performance is to 'feed forward' (as opposed to feedback) focusing on what students should concentrate on in future tasks[10]

Summative Assessment

Summative assessments are intended to identify how much the student has learnt. CBDs have a strong summative feel: a minimum number of assessments are required and a satisfactory standard must be reached to allow progression of a trainee to the next level of training[1]. Summative assessment affects students in a number of different ways; it guides their judgment of what is important to learn, affects their motivation and self-perceptions of competence, structures their approaches to and timing of personal study, consolidates learning, and affects the development of enduring learning strategies and skills[12-13]. Resnick and Resnick summarize this as "what is not assessed tends to disappear from the curriculum"[13]. Accurate recording of CBDs is vital, as the assessment process is transient, and allows external validation and moderation.

Evaluation of any teaching is fundamental to ensure that the curriculum is reaching its objectives[14]. Student evaluation allows the curriculum to develop and can result in benefits to both students and patients. Kirkpatrick suggested four levels on which to focus evaluation[14]:

Level 1 – Learner's reactions
Level 2a – Modification of attitudes and perceptions
Level 2b – Acquisition of knowledge and skills
Level 3 – Change in behaviour
Level 4a – Change in organizational practice
Level 4b – Benefits to patients

At present there is little opportunity within the Intercollegiate Surgical Curriculum Project (ISCP) for students to provide feedback. Thus a typical 'evaluation cycle' for course development (figure 2) cannot take place[15]. Given the widespread nature of subjects covered by CBDs, the variations in marking standards by assessors, and concerns with validity and reliability, an overall evaluation of the curriculum may not be possible.
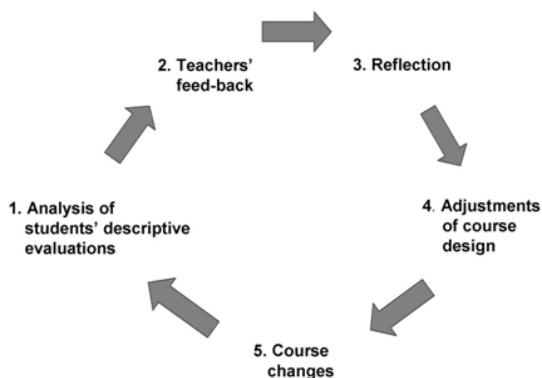

Figure 2: Evaluation cycle used to improve a teaching course[15]

However, regular evaluation of the learning process can improve the curriculum and may lead to better student engagement with the assessment process[14]. Ideally the evaluation process should be reliable, valid and inexpensive[15]. A number of evaluation methods exist, but all should allow for ongoing monitoring review and further enquiries to be undertaken.

Conclusion

CBDs, like all assessments, do have limitations, but we feel that they play a vital role in development of trainees. Unfortunately, Pereira and Dean suggest that trainees view CBDs with suspicion[7]. As a result, students do not engage fully with the assessment and evaluation process and CBDs are not being used to their full potential. The main problems with CBDs relate to the lack of formal assessor training in the use of the WBA and the lack of evaluation of the assessment process Adequate training of assessors will improve feedback and standardize the assessment process nationally. Evaluation of CBDs should improve the validity of the learning tool, enhancing the training curriculum and encouraging engagement of trainees.

If used appropriately, CBDs are valid, reliable and provide excellent feedback which is effective and efficient in changing practice. However, a combination of assessment modalities should be utilized to ensure that surgical trainees are facilitated in their development across the whole spectrum of the curriculum.

Competing Interests
None declared
Author Details
J M L Williamson, MBChB, MSc, MRCS, Specialty Training Registrar. A J Osborne MBBS, MRCS, Specialist Registrar, Department of Surgery, The Great Western Hospital, Marlborough Road, Swindon, SN3 6BB
CORRESSPONDENCE: J M L Williamson MBChB, MSc, MRCS. Specialty Training Registrar, Department of Surgery, The Great Western Hospital, Marlborough Road, Swindon, SN3 6BB
Email: jmlw@doctors.org.uk

REFERENCES

1. Intercollegiate Surgical Curriculum Project (ISCP). ISCP/GMP Blueprint, version 2. ISCP website (www.iscp.ac.uk) (accessed November 2010)

2. Lake FR, Ryan G. Teaching on the run tips 4: teaching with patients. Medical Journal of Australia 2004;181:158-159

3. Swanson DB, Norman GR, Linn RL. Educational researcher 1995;24;5-11+35

4. Miller GE. The assessment of clinical skills/competence/ performance. Academic Medicine, 1990;65:563–567.

5. Feldt LS, Brennan RL. Reliability. In Linn RL (ed), Education measurement (3rd edition). Washington, DC: The American Council on Education and the National Council on Measurement in Education; 1989

6. Moss PA. Can there be Validity without Reliability? Educational Researcher 1994;23;5-12

7. Pereira EA, Dean BJ. British surgeons' experience of mandatory online workplaced-based assessment. Journal of the Royal Society of Medicine 2009;102:287-93

8. Veloski J, Boex JR, Grasberger MJ, Evans A, Wolfson DB. Systemic review of the literature on assessment, feedback and physician's clinical performance: BEME Guide No. 7. Medical teacher 2006;28:117-28

9. Norcini J, Burch V. Workplaced-based assessment as an educational tool: AMEE Guide No 31. Medical teacher 2007;28:117-28

10. Hounsell D. Student feedback, learning and development in Slowery, M and Watson, D (eds). Higher education and the lifecourse. Buckingham; Open University Press; 2003.

11. Bloxham S, Boyd P. Developing effective assessment in higher education: A practical guide. Maidenhead: Open University Press; 2007

12. Cooks TJ. The impact of classroom evaluation practices on students. Review of Educational Research 1998;58;438-481

13. Resnick LB, Resnick D. Assessing the thinking curriculum: New tools for educational reform. In Gifford B and O'Connor MC (eds), Cognitive approaches to assessment. Boston: Kluwer-Nijhoff; 1992

14. Barr H, Freeth D, Hammick M, Koppel, Reeves S. Evaluation of interprofessional education: a United Kingdom review of health and social care. London: CAIPE/BERA; 2000

15. Wahlqvist M, Skott A, Bjorkelund C, Dahlgren G, Lonka K, Mattsson B. Impact of medical students' descriptive evaluations on long-term course development. BMC Medical Education 2006;6:24