The American Journal of Surgery 193 (2007) 86–89
Association for Surgical Education

Is it feasible to include a technical skill station on a national licensing examination?

Martin Friedlich, M.D., M.Sc., M.Ed., F.R.C.S.C.[a,*], Sydney Smee, M.Ed.[b], Timothy Wood, Ph.D.[b]

[a]*Department of Surgery, University of Ottawa, Ottawa, Ontario, Canada*
[b]*Medical Council of Canada*

Manuscript received April 7, 2006; revised manuscript September 13, 2006

Abstract

The purpose of the current study was to assess the feasibility and validity of including a technical skill station on a national licensing examination. At the 2003 Medical Council of Canada Qualifying Examination, 745 test takers participated in a pilot station assessing the ability to perform a technical procedure. Checklists and rating scales were used for scoring. Validity was investigated by comparing surgery-trained to non–surgery-trained test takers. The mean for the pilot station was 72.4%. The pilot station was moderatley correlated to the rest of the examination (item-total correlation .43). The mean score for surgery test takers was higher than for other test takers ($P < .001$). Inclusion of a technical skill station on a high-stakes examination is feasible, and at many levels, there is evidence of the validity of including this station.

Keywords: Technical; Skill; Evaluation; National; Examination

Although the use of technical skills for the performance of minor surgical procedures is common in primary care, little assessment of these skills actually occurs during training. Furthermore, these skills are not assessed on either national licensing examinations or on certification examinations. This failure to assess technical skills at the high-stakes level may occur because the importance of these skills are not recognized, because it is assumed that these skills are assessed in-training or because the ability to test for these skills in a cost-effective and reliable manner has not been possible.

Technical skill evaluation has been studied in certain medical specialties. For example, the Objective Structured Assessment of Technical Skill (OSATS) was developed for the appraisal of technical skill in surgical trainees [1,2]. The OSATS platform consisted of surgical trainees rotating amongst skill stations, while being evaluated by surgeons using checklists and global rating scales. This examination is statistically reliable and satisfies many important validity issues [3]. A similar evaluation tool has been developed for the assessment of technical skill in family practice residents [4]. The Structured Assessment of Minor Surgical Skill (SAMSS) also has moderate to high reliability and shows evidence of construct validity. The SAMSS format has also been used to assess the technical skills of clinical clerks [5]. This assessment of clinical clerks proved feasible and demonstrated similar reliability as achieved with the assessment of family medicine residents. This finding suggests that the assessment of technical skills could be implemented during medical training.

What about high-stakes examinations? Most licensing and certifications examinations typically test candidates' history taking, physical examination skills, clinical judgment, and communication skills. If the assessment of technical skills were to be implemented on a high-stakes examination, it would be necessary both to study the feasibility of assessing these skills and to determine how the performance on a technical skills station related to performance on the rest of the examination. A recent study in the anesthesia literature compared a simulator-based examination with an oral examination for assessing the management skills of senior anesthesia residents. Although the oral examination was modeled on a genuine board certification examination, neither the simulator-based examination nor the oral examination was actually applied at the certification level [6]. The Medical Council of Canada decided to investigate the possibility of using a technical skills

* Corresponding author. Tel.: +1-613-739-6650; fax: +1-613-739-6646.
E-mail address: mfriedlich@ottawahospital.on.ca

0002-9610/07/$ – see front matter
doi:10.1016/j.amjsurg.2006.09.006

station. Candidate performance on the pilot station was assessed and compared with performance on the other stations in the examination. The construct validity of the station was assessed by comparing the scores of participants who were trained in surgery related postgraduate training programs to participants trained in other types of postgraduate training programs. Those with a surgical background should do better on a station assessing technical skill, and therefore a better performance by surgical residents than nonsurgical residents would provide evidence of construct validity. The work involved in creating such a station was also tracked.

Methods

Participants

A total of 1911 test takers attempted the November 2003 Medical Council of Canada Qualifying Examination Part II (MCCQE Part II), a prerequisite for medical licensure in Canada. All test takers have successfully completed a minimum of 12 months of postgraduate training and must have passed the Medical Council's Qualifying Examination Part I (MCCQE Part I), an assessment of knowledge and clinical reasoning commonly taken at the end of medical school. The pilot station from this study was administered at 5 of the 15 possible examination centers and a total of 745 (39%) test takers participated. The 5 sites were chosen to be geographically diverse and all were English-speaking locations.

Design

The MCCQE Part II uses physician examiners and standardized patients (SPs) to assess communication skills, problem solving skills, data acquisition skills, and legal, ethical and organizational issues. The content of each station is derived from 1 of the 5 major disciplines of medicine (medicine, obstetrics and gynecology, pediatrics, psychiatry, and surgery) and each station is linked to one of the MCC Objectives, which are used to ensure content validity of the MCC examinations [7]. The stations are designed to test skills and judgment common to all (undifferentiated) physicians and must consist of problems that are common and/or critical for the practice of medicine.

The examination itself is a 12-station OSCE consisting of six 10-minute patient encounters and 6 couplets stations. The couplets are 5-minute patient encounters paired with 5-minute paper and pencil exercises. In addition to these 12 stations, the examination also includes 2 pilot stations. These stations, which consist of novel content, are not included in the scoring. Within an examination center, 1 to 4 different tracks are run depending on the number of candidates at that center. Each candidate has a different starting station within their track and everyone circulates until they have completed all of the stations, so the order of stations is different for each candidate. Candidate instructions, including a clinical vignette and the task that is to be performed, are located outside each station.

Within each station, an examiner directly observes each candidate and for each one completes a checklist of the relevant clinical tasks, a rating scale that assesses the quality of the patient interaction, and rates each one on a 6-point global item. The data from this last item are used in the modified borderline method to set the cut score for the station [8,9].

For this study, candidates were asked to demonstrate their ability to perform a basic minor surgical procedure. The examiners scored both a checklist and a 4-item rating scale assessing respect for tissue, time and motion, technical ability, and universal precautions. The scores from the checklist and the rating scale were each worth 50% of the station score. To respect the security of examination content, the specific procedure is not described here; however, the procedure is a basic technical procedure commonly performed by primary care physicians [10–13] and selected by the MCCQE Part II examination test committee. This station was treated as a pilot station on the examination. Candidates are aware that 2 of the stations on the examination are pilot stations and will not count toward their mark; however, they are not told which 2 stations are the pilot stations.

Forty-eight surgeons and emergency physicians were selected as examiners to ensure that all of those scoring in this station performed minor surgical procedures regularly. Only 1 evaluator was used per station. The candidates are blinded to which stations are pilots, so double examiners would compromise this standard. In addition, stations are run in small clinic rooms, which makes for a tight fit with just 1 examiner. Furthermore, examiners are not only difficult to find but also expensive so this study strived for a maximum sample of candidates and relied on the results being averaged across a large number of examiners.

Ethics

This study met the research guidelines that were in place at the Medical Council of Canada in 2003. These guidelines included reviews by the MCC OSCE Test Committee, which oversees examination content, and the MCC Central Examination Committee, which oversees operation of all MCC examinations. Throughout the review process consideration is given not only to the relevance of the content but also to the appropriateness of the presenting problem and task given the time limit imposed by the OSCE design and to the impact a new format or patient affect may have on candidates. No difficulties were anticipated for this station.

Analysis

The purpose of this study was to assess the feasibility and validity of including a technical skills station on a high-stakes clinical examination. Two analyses were run. First, the feasibility was analyzed by comparing scores on the technical station to scores on the other stations on the examination. Second, the study investigated the validity of the technical skills station by comparing surgery-trained test takers to non–surgery-trained test takers. For this study, test takers from the following postgraduate training programs were classified as surgery-trained test takers: cardiac surgery, general surgery, neurosurgery, ophthalmology, orthopedic surgery, otolaryngology, plastic surgery, urology, and obstetrics/gynecology. Test takers from all other types of postgraduate training programs were classified as non–surgery test takers.