RESEARCH ARTICLES

Objective Structured Clinical Examinations in Doctor of Pharmacy Programs in the United States

Deborah A. Sturpe, PharmD

University of Maryland School of Pharmacy

Submitted October 9, 2009; accepted June 11, 2010; published October 11, 2010.

**Objectives.** To describe current objective structured clinical examination (OSCE) practices in doctor of pharmacy (PharmD) programs in the United States.

**Methods.** Structured interviews were conducted with PharmD faculty members between September 2008 and May 2010 to collect information about awareness of and interest in OSCE, current OSCE practices, and barriers to OSCEs.

**Results.** Of 108 US colleges and schools of pharmacy identified, interviews were completed for a representative sample of 88 programs (81.5% participation rate). Thirty-two pharmacy programs reported using OSCEs; however, practices within these programs varied. Eleven of the programs consistently administered examinations of 3 or more stations, required all students to complete the same scenario(s), and had processes in place to ensure consistency of standardized patients’ role portrayal. Of the 55 programs not using OSCEs, approximately half were interested in using the technique. Common barriers to OSCE implementation or expansion were cost and faculty members’ workloads.

**Conclusions.** There is wide interest in using OSCEs within pharmacy education. However, few colleges and schools of pharmacy conduct OSCEs in an optimal manner, and most do not adhere to best practices in OSCE construction and administration.

**Keywords:** objective structured clinical examination (OSCE), assessment, testing, examination

INTRODUCE

The OSCE was introduced by Dr. Ronald M. Harden in the 1970s as “an approach to the assessment of clinical competence in which the components of competence are assessed in a planned or structured way with the attention being paid to the objectivity of the examination.”1 The examination consists of multiple, standard stations at which students must complete 1 to 2 specific clinical tasks, often in an interactive environment involving patient actors (ie, standardized patients).1,2 OSCE has become a common method to test learner performance across a variety of health professions disciplines. Most notably, OSCE is a component of entry-to-practice licensing examinations, including the United States Medical Licensing Examination, the Canadian Pharmacist Qualifying Examination, and the Medical Council of Canada Qualifying Examination.3-5 Interest in the OSCE technique appears to be growing within US colleges and schools of pharmacy, as evidenced by a 7-fold increase in OSCE research abstracts presented at national academic pharmacy meetings between 2006 and 2009.6-9

As use of OSCE grows within pharmacy education, it is important to ascertain whether the technique is applied in a way that maintains examination reliability and validity, especially if colleges and schools of pharmacy plan to use the technique as a component of high-stakes assessments. Although individual colleges ultimately must determine whether such metrics are met, the works of Harden1 and others provide some recommendations for general examination procedures and techniques that should contribute positively to these measures. However, no definitive “standard” is available in the literature that defines minimally acceptable practices.

To maintain examination validity and authenticity, a representative sampling of real-world skills should be tested.10 Consequently, use of a blueprint that defines examination domains (eg, knowledge, skills, behaviors, complexity) to guide OSCE station development along with group (rather than individual) writing of OSCE cases with peer review has been recommended.2,10-12 Maintaining station consistency through careful standardized patient and examiner (the person who scores student performance)
training and quality assurance processes on examination day, pilot testing of stations, and establishment of objective pass/fail guidelines should also increase examination reliability. Inclusion of an appropriate number of examination stations is an important consideration to reduce sampling error. Harden suggests that “the reliability of the examination is, in a large measure, dependent on the number of independent assessments of competence made during the examination.” Although no single definition is available to determine what this optimal number of independent assessments may be, OSCEs used for medical and pharmacy licensing purposes in the United States and Canada range from 12 to 16 stations (with a small percentage of those stations being piloted stations only). Harden’s original OSCE was comprised of 16 stations. Less information is available regarding the minimum number of stations required for non-high-stakes summative OSCEs, although one could reasonably assume that more than 1 or 2 independent assessments should be made in such cases, based on Harden’s original concept that the examinations should consist of multiple stations. A summary of these general recommendations as well as some additional best practices suggested by Harden are provided in Table 1.

Little is known about how OSCEs are implemented within US colleges and schools of pharmacy. Consequently, the goal of this study was to describe current OSCE practices in the United States.

**METHODS**

A structured interview instrument was developed based on the literature, personal experience with OSCEs, and consultations with 3 pharmacy faculty members in the United States and Canada, who had experience using OSCEs for either programmatic or pharmacy licensing examination purposes. The structured interview technique was chosen over other survey techniques because it is more flexible and adaptive, and allows for natural conversation, enabling the researcher to adjust the direction of the interview based on previous responses. The final interview instrument consisted of 3 content areas for programs not using OSCEs (introductions and definitions, interest in and barriers to OSCEs, and program demographics), and 7 content areas for programs using OSCEs (introductions and definitions, OSCE’s role in the curriculum, OSCE development process, standardized patient information, OSCE scoring process, OSCE logistics, and program demographics). All US PharmD programs with either precandidate, candidate, or full accreditation status as listed on the Accreditation Council for Pharmacy Education (ACPE) Web site in August 2008 were eligible for inclusion in the study. Colleges and schools of pharmacy were contacted initially by the author on a rolling basis by e-mail between August 2008 and February 2009 to invite participation. This initial contact was made with the person responsible for assessment or curriculum, if such an individual could be identified on the program’s Web site; otherwise, initial contact was made with the chair of the pharmacy practice department. The e-mail explained the purpose of the project and asked for referral to the faculty member most appropriate to contact regarding OSCEs, with a follow-up e-mail sent to that person if identified. Once responses to these initial contact e-mails were received, individual appointments for a structured interview were scheduled with the appropriate faculty member(s) at each institution.

For colleges and schools of pharmacy that did not respond, follow-up e-mails were sent within 3 to 6 months of the initial contact to reinvite participation, and some programs were selectively contacted by telephone if known to use OSCE, if the author personally knew a faculty member at the college or school of pharmacy, or if located in a geographic area that was underrepresented to that point in the study. To attempt to maximize the response rate and achieve greater than 80% participation from colleges and schools of pharmacy, final e-mail and telephone contacts were made between January and April 2010.

All interviews were conducted by the author between September 2008 and May 2010. Interviews required approximately 15 to 20 minutes for colleges and schools of pharmacy not using OSCEs and 60 minutes for programs using OSCEs. The author recorded notes during the interview directly on the interview instrument, then immediately transferred those notes into an electronic format once the interview was completed. Interviews were not audiotaped. Descriptive statistics were used to analyze interview transcripts.

For the purpose of this project, the following definitions were used: (1) “OSCE” was defined as an

<table>
<thead>
<tr>
<th>Table 1. Recommended Practices to Improve Objective Structured Clinical Examination(OSCE) Validity and Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blueprint guides station and examination development</td>
</tr>
<tr>
<td>Group casewriting</td>
</tr>
<tr>
<td>Case peer review (“validation”)</td>
</tr>
<tr>
<td>Establish objective pass/fail guidelines</td>
</tr>
<tr>
<td>All students complete the same task(s)</td>
</tr>
<tr>
<td>Train standardized patients and examiners prior to the</td>
</tr>
<tr>
<td>examination</td>
</tr>
<tr>
<td>Pilot test stations</td>
</tr>
<tr>
<td>Perform examination day quality assurance</td>
</tr>
<tr>
<td>Pick the right number of stations to test wide range of skills</td>
</tr>
<tr>
<td>(up to 12-16 stations for high-stakes examinations)</td>
</tr>
</tbody>
</table>
examination that is given for the purpose of assessing student performance on clinical tasks, and designed so that examination components are standardized (all students experience the same scenario) and objective (performance expectations are set prior to the examination); (2) “Summative” was defined as an assessment in which performance contributed to a student’s course grade or determine progression in the program; (3) “Formative” was defined as an assessment in which performance did not contribute to a student’s course grade or determine progression in the program (eg, the assessment was for feedback purposes only, or all students received participation points); (4) “High stakes” was defined as an assessment in which poor performance on the assessment could prevent progression in the pharmacy program (eg, course failure); and (5) “Low stakes” was defined as an assessment that could contribute to a student’s grade, but would not prevent progression through the PharmD program if poor performance was observed. Because the nature of this project was to collect information about colleges and schools of pharmacy and not about living individuals, the Institutional Review Board at the University of Maryland, Baltimore determined that the study did not require review.

RESULTS

One hundred eight colleges and schools of pharmacy were invited to participate in this study; 88 interviews were conducted for a participation rate of 81.5%. Participation rates by subgroups of geographic region, accreditation status, and institution type is shown in Table 2. During the study, 1 electronic interview record was corrupted and determined to be nonrecoverable; thus, 87 program responses were included in the analysis. Of those 87 programs, 32 (37%) reported using OSCEs, and 55 (63%) reported that they did not use OSCEs in their curriculum. A demographic comparison of programs using and not using OSCEs is presented in Table 3.

All but 2 of the 55 colleges and schools of pharmacy not using OSCEs were aware of the technique, and approximately half were considering incorporating OSCEs into curricular activities within the next few years, although just 20% reported being in active planning for implementation. Barriers identified to implementing OSCEs included cost (n = 34); concerns over increased faculty workloads (n = 29); lack of faculty awareness or buy-in to the technique (n = 14); lack of access to a standardized patient program (n = 13); concerns over the validity and reliability of the technique compared to other assessment methods (n = 11); difficulty incorporating OSCEs into an existing curriculum (n = 9); and lack of space to conduct OSCE activities (n = 9). Twelve of the 55 programs reported hiring standardized patients for teaching and assessment purposes within the curriculum. Reasons cited why these activities were not OSCEs included lack of a multistation examination, ie, only 1 to 2 stations (n = 9), inability to ensure consistency of the standardized patient role portrayal (n = 6), and that not all students encountered the same scenario(s) (n = 5).
Of the 32 colleges and schools of pharmacy using OSCEs within the curriculum, most reported implementing OSCEs due to interest by faculty members to use a technique better suited to assessing integration of student knowledge, skills, and communication, compared to traditional methods of assessment. Eleven reported implementing OSCEs in direct response to the ACPE Standards 2007. Fourteen (44%) reported that at least 1 of their faculty members had attended some sort of training that specifically focused on the OSCE technique (often a multi-day conference).

Twenty-one colleges and schools of pharmacy reported using OSCEs as an assessment technique within courses, 4 used OSCEs as a standalone programmatic assessment tool, and 7 used OSCEs for both purposes. For colleges and schools of pharmacy placing OSCEs within courses, the most common course types included integrated laboratory courses (n = 13), communications courses (n = 6), pharmacotherapeutics courses (n = 4), and advanced pharmacy practice experiences (APPE) (n = 4). Almost all (n = 30) used OSCEs for summative assessment purposes, with 18 also using OSCEs for formative assessments. Eight programs conducted high-stakes OSCEs. Most not yet using OSCEs for high-stakes purposes reported interest in progressing towards this goal, often for developing standalone capstone examinations that could be administered prior to the start of APPEs and graduation.

The 18 colleges and schools of pharmacy using OSCEs for formative purposes provided students with their raw scores on the examinations. Additionally, 5 of these programs required students to meet with the standardized patient immediately following the encounter for verbal performance feedback; 6 had students meet with faculty members after the encounter for verbal performance feedback; 2 provided students with copies of their scoring rubric; and 1 provided general written feedback to students about cohort (but not individual) performance. Ten colleges and schools required students to view their performance on video tape after OSCE encounters.

Of the 30 colleges and schools of pharmacy using OSCEs for summative purposes, 6 provided no feedback to students about their performance, and 10 provided raw performance scores only. Students were less likely to receive individual feedback from standardized patients (n = 3) or faculty members (n = 2), although individual feedback was provided by 4 colleges and schools if remediation was required after OSCE failure.

The number of stations contained within any single OSCE varied greatly across colleges and schools of pharmacy and within program, depending on the OSCE’s purpose. Some administered OSCEs at 1 station only, while others administered OSCEs at ≥ 8 stations. Thirty-one percent (n = 10) purposefully designed OSCEs so that each student encountered different scenario(s). Station selection was guided by blueprint development in 28% of colleges and schools (n = 9), and stations were pilot tested in just 19% (n = 6). Few programs used group caseworking methods to develop OSCE stations (31%, n = 10), and just 47% (n = 15) had stations validated through a review process.

Sixty-three percent (n = 20) of colleges and schools established an absolute pass/fail standard for each OSCE station, while 37% (n = 12) awarded points per item on OSCE checklists without setting a minimum passing standard. Of the 20 colleges and schools setting a pass/fail standard, 4 used the Angoff method,15 1 used the Borderline method,16 and the remainder used an arbitrary method (eg, a random pass point of 70% was chosen).

A variety of methods were reported by colleges and schools to maintain examination security, including not publishing or providing scoring rubrics to students (n = 11); purposeful use of different cases within the same examination (n = 6); sequestration of students on the day of the examination (n = 5); requiring students to sign a confidentiality agreement (n = 4); and verbally reminding students that the program’s honor code prohibited discussing examination content (n = 3).

The characteristics of the person who served in the standardized patient role varied across and within PharmD programs, with several using a variety of standardized patient types. In total, 20 colleges and schools hired professional standardized patients for some or all of their OSCE activities, while many relied on using pharmacy faculty members (n = 6), pharmacy residents (n = 4), pharmacy students (n = 7), and non-pharmacy volunteers, such as administrative staff or spouses (n = 5). For those hiring professional standardized patients and able to report salary data, salaries ranged from $13 to $25 per hour. This cost usually was covered by an administrative budget (ie, dean’s office or department budget), but 5 colleges and schools of pharmacy did report paying for OSCEs through student fees. While 63% (n = 20) trained standardized patients prior to the examination day, 25% (n = 8) admitted to providing minimal training, same-day training, or no training at all. Data regarding how standardized patients were trained was not known by the persons being interviewed for 4 of the PharmD programs included in the study.

Forty-seven percent (n = 15) of the colleges and schools reported that the standardized patient also served as the examiner during OSCEs, while the other programs reported using pharmacy faculty members or residents. For those using this secondary examiner, a third provided
examiner training prior to the examination day, with the remainder providing same-day training.

Half of the colleges and schools of pharmacy interviewed reported the ability to provide quality assurance for consistent role portrayal by the standardized patient on examination day. Most accomplished this by asking pharmacy faculty members and/or standardized patients to watch encounters on remote video monitors and report deviations to the examination coordinators. Fifty-nine percent (n = 19) of PharmD programs were able to record (with videotape or digital methods) student encounters for later viewing if performance was challenged.

Of the 8 colleges and schools of pharmacy using OSCEs for high-stakes purposes, 5 incorporated formative and low-stakes summative OSCEs in the curriculum prior to the high-stakes examination to provide students with the opportunity to practice being tested and evaluated using the technique; but 3 PharmD programs administered the high-stakes examination as the first OSCE experience. Four colleges and schools used a blueprint to guide station development, and 5 validated cases. Five programs used an arbitrary standards-setting method; 1 purposefully tested students on differing skills; 5 did not pilot test cases; and 4 were not able to ensure consistent role portrayal by the standardized patient. Only 3 PharmD programs administered an OSCE with ≥ 5 stations, while 2 administered an OSCE of only 1 or 2 stations.

Methods varied for OSCE remediation if failure occurred. Four of the 8 colleges and schools used OSCEs as a high-stakes examination prior to graduation. Of these, 2 administered the examination in December to allow for remediation activities during the spring semester (such as rescheduling remaining APPEs); 1 then retested students immediately prior to graduation while the other did not, with the former delaying graduation if the retest was not passed. Another PharmD program extended the examination from 3 to 6 stations if students failed the initial 3-station examination, to increase the breadth of skill sampling, with students not passing the expanded examination having graduation delayed. The final PharmD program did not have a remediation plan in place and reported no OSCE failures since the examination was implemented. Of the 4 using OSCEs as a high-stakes examination within a course, 1 allowed students to retake the examination an unlimited number of times until they passed, 2 provided feedback to students and then allowed a single retest with subsequent course failure if not passed, and 1 did not have a remediation plan in place and reported no OSCE failures since the examination was implemented.

Just 34% (n = 11) of colleges and schools using OSCEs consistently administered examinations with 3 or more stations, required all students to complete the same scenarios, and ensured consistency of standardized patient portrayal. Of the 8 PharmD programs using OSCEs for high-stakes purposes, only 3 met these criteria.

DISCUSSION

This study provides important insight regarding use and quality of OSCEs in PharmD curricula in the United States. Many colleges and schools of pharmacy either use OSCEs or are interested in starting OSCEs to measure clinical competence accurately, but barriers to OSCE implementation and expansion relate to cost and manpower issues. Survey results of standardized patient programs across the United States and Canada revealed that the average standardized patient salary was $16 per hour, with a range of $10 to $30 per hour, similar to salaries reported in this study. This estimate does not account for additional costs such as space, administrative overhead, and faculty members' time. Unfortunately, a paucity of published data exists that reports the total cost of OSCEs, but estimates have ranged from $21 to $1000 per examinee, depending on the total number of stations. Also, scant data exists regarding total manpower hours required to develop and administer OSCEs, although Cusimano and colleagues approximated that 8.2 person-hours per student were required to develop and implement a 6-station OSCE (1.4 hours per student per station).

Perhaps partly due to these financial and time constraints, many US colleges and schools of pharmacy have developed OSCEs with a low number of stations, use individuals instead of groups to write cases, use nonprofessional standardized patients, and have limited ability to ensure consistency of the patient's role portrayal either due to limited training and/or lack of space and technology that allows viewing of encounters. In such cases, it is questionable whether or not an OSCE as defined by Harden is being administered. Twelve PharmD programs hired standardized patients in teaching and assessment activities, yet recognized they were not performing an OSCE due to the low number of stations used, the inability to ensure consistency of patient role portrayal, and variance of the scenarios tested. Other PharmD programs reported conducting OSCEs despite having these same characteristics. This confusion suggests the need to publicize more effectively a standard definition of OSCE so that it may be distinguished from other types of performance-based assessments.

If colleges and schools of pharmacy are to move toward implementation of high-stakes OSCEs for assessment of student performance at critical stages of development (such as at the conclusion of educational training prior to starting APPEs, or prior to graduation to validate overall quality of the educational and experiential
program), PharmD programs must create reliable and validated examinations. Simple changes to present practices that may help elevate OSCE quality include consistent use of blueprinting, training of standardized patients and examiners prior to examination day, and use of a validated method to establish passing standards. Creation of authentic and realistic cases through group casewriting and validation remains a challenge. One solution to this problem would be the establishment of regional consortiums that collectively write and validate cases. Such consortiums have been established in California and through the “Big 10” colleges and schools of pharmacy. Finally, the most difficult change to implement is an increase in the number of OSCE stations within an examination so that a reasonable scope of knowledge and skills is tested. As previously mentioned, OSCEs used for medical and pharmacy licensing purposes in the United States and Canada range from 12 to 16 stations, a number that is higher than reported by most PharmD programs in this study, including those using OSCEs for high-stakes purposes. Consequently, colleges and schools not able to develop and sustain OSCEs of this scope may need to concentrate on using OSCEs primarily for low-stakes summative and/or formative purposes.

The primary limitation to this research was the data collection technique itself. A single investigator with no previous training or experience in conducting structured interviews collected all data. For this reason, there is no guarantee that each interview was conducted with equivalent vigor and attention to detail, especially because the interviews were completed over 2 academic years. There is also no guarantee that all responses by the interviewee were recorded and reported accurately by the investigator, because interviews were not audiotaped. However, because a single investigator completed all interviews using a specific and detailed interview tool, and because the same process was used for each interview, reasonable standardization did exist during the data collection and reporting process. Because the intention of the research was to report on current trends in OSCE use in US colleges and schools of pharmacy, these limitations were not great enough to invalidate the study’s findings.

CONCLUSIONS

Despite widespread interest in using OSCEs within pharmacy education, few PharmD programs in the United States are conducting OSCEs in a reliable and valid manner. Key best practices for OSCE construction and administration are not implemented consistently in many programs, including programs using OSCEs for high-stakes purposes.

ACKNOWLEDGMENTS

The author wishes to acknowledge the following persons for their assistance in developing the structured interview instrument: Zubin Austin, PhD; Mary Beth O’Connell, PharmD; and Francine Salinitri, PharmD.

REFERENCES