The Role of Standardized Patient and Trainer Training in Quality Assurance for a High-stakes Clinical Skills Examination

Gail E. Furman
Clinical Skills Evaluation Collaboration, National Board of Medical Examiners, Philadelphia, USA.

For over 30 years, medical educators have used standardized patients (SPs), laypersons trained to portray a patient case in a realistic manner, to teach and to assess clinical skills. All medical schools in the US have SP programs in place, and the US and Canada require national examinations using SPs to assess the competency of those wishing to obtain licensure to practice medicine in these countries. To ensure a valid and reliable examination, unwanted variance that can be introduced by SP performance must be addressed. The goal of SP training is to imbue the SP with the characteristics, mannerisms and history of a real patient so that the portrayal is consistent and accurate. The challenge is to ensure consistent portrayal of each case with sufficient realism to elicit the expected clinical performance and to ensure standardized SP performance across multiple examinees. This paper considers the quality assurance methods applied to training the SP trainers and the protocols used to train the SPs, to ensure that the SP performances are sufficiently accurate and standardized, and that the evaluators completing the checklists and scales used for scoring do so correctly and consistently.

Key Words: clinical skills examinations, licensure examination, standardized patients
(*Kaohsiung J Med Sci* 2008;24:651–5)

For over 30 years, medical educators have used standardized patients (SPs), laypersons trained to portray a patient case in a realistic manner, to teach and to assess clinical skills. All medical schools in the US have SP programs in place. The US and Canada require national examinations using SPs to assess the competency of those wishing to obtain licensure to practice medicine in these countries.

Received: Jan 5, 2009 Accepted: Jan 22, 2009
Address correspondence and reprint requests to: Dr Gail E. Furman, Director of Quality Assurance, CSEC, National Board of Medical Examiners, 3750 Market Street, Philadelphia, PA 19147, USA.
E-mail: gfurman@nbme.org

The USMLE Step 2 CS

The United States Medical Licensing Examination™ (USMLE™) is a three-step examination for medical licensure in the US and is sponsored by the Federation of State Medical Boards and the National Board of Medical Examiners® (NBME®). The Step 2 Clinical Skills (CS) examination assesses whether examinees can demonstrate the fundamental clinical skills essential for safe and effective patient care under supervision. Step 2 CS uses SPs, and the cases cover common and important situations that a physician is likely to encounter in common medical practice in clinics, doctors' offices, emergency departments and hospital settings in the US.

Examinees have 15 minutes to interact with the SP in an examination room, where they are expected to

© 2008 Elsevier. All rights reserved.

establish rapport with the patients, elicit pertinent historical information from them, perform focused physical examinations, answer questions, and provide counseling as appropriate. After each interaction with a patient, examinees have 10 minutes to record pertinent history and physical examination findings, list diagnostic impressions, and outline plans for further evaluation, if necessary.

SPs record the examinee's history taking questions and physical examination maneuvers using checklists, and record the communication, interpersonal skills, and spoken English proficiency of the examinee using rating scales. The physician raters score the completed patient notes. The checklists and patient note scores make up the Integrated Clinical Encounter scores; separate scores are reported for communication and interpersonal skills (CIS) and spoken English proficiency (SEP).

This paper will summarize some of the methods used by the Clinical Skills Evaluation Collaboration (CSEC) to ensure standardization of SP training and quality assurance checks on performance during the USMLE Step 2 CS examination.

Using SPs to Assess Clinical Skills

There are several reasons for using SPs in assessing clinical skills. For the USMLE and Medical Council of Canada exams, SPs are used as a proxy for a "real patient" to observe how the examinee will perform with respect to history taking and physical examination skills. In the US, the SPs also record interpersonal and communication skills, as well as spoken English proficiency. SPs cost less to train and to employ compared with the cost of employing academic physicians to score the exam. For an examination that takes place throughout an entire year, SPs are also easier to schedule as compared with the schedules of academic physicians. Providing intelligent SPs are hired, the SPs can be easy to train and will follow directions well. The main reason for using SPs is standardization, especially if multiple sites are used for testing. Many conditions can be simulated, so several people can be recruited and trained to portray the same case. A summary of some of the physical findings that can be portrayed by standardized patients has been reported elsewhere, and includes neurological symptoms, acute and chronic pain, and shortness of breath, for example [1].

Studies have shown that SPs are realistic in the manner in which a "real" patient is portrayed [2]. One study sent unannounced SPs into physician's offices, and doctors were unable to identify these patients as "simulated" [3]. Another study recorded how physicians interacted with SPs and found that the doctors performed in a similar manner with respect to history taking and physical examination of real patients in the clinical setting [4].

In addition to being realistic, SPs are accurate in the portrayal of simulated cases and in documenting the actions of the examinee. On average, SPs are more than 90% accurate in portraying the details of the case they are trained to represent [5]. One study showed that, with limited training, laypersons can reliably evaluate 83% of the same clinical skills that were evaluated by faculty physicians. With more training, SPs can approach being 100% accurate. Some skills are easier to experience than to observe; for example, abdominal palpation (how deep, how much pressure). Moreover, other studies have shown that several SPs portraying the same case do not significantly impact the reliability of the exam [5].

Hiring SPs

The background of the SPs used in the USMLE Step 2 CS examination varies. SPs are actors, housewives, teachers, and retired people. The current SP population is between 18–75 years of age, with sex equally represented. To characterize the cultural diversity of the US patient population, SPs are African-American, Asian, Hispanic, and Caucasian. Those interested in becoming SPs use a company website to apply online to start the enrolment process.

One study using the NEO-PIR personality inventory showed that SPs currently working in the USMLE Step 2 CS examination are psychologically normal people, whose personality traits were statistically significant for extroversion and openness to experiencing new things [6].

Standardization of SPs

To deliver valid and reliable examinations for the purpose of licensure, one goal is to eliminate potential site differences. To achieve this goal, all trainers follow

the same training protocol for training at each site; all SPs are trained to respond in a standardized way, no matter which of the five testing centers an examinee attends. A web-based program developed by the NBME called eCase, is used to train all SPs. The eCase is a multimedia approach to training, which includes interactive quizzes, and links to demonstration videos and encounters. All case material is stored in a secure server, which can be accessed at each site by the trainers and SPs. However, for security reasons, staff and SPs can only access cases they are authorized to see. All cases requiring a physical examination are linked to a standardized physical examination database, which contains videotapes showing which maneuvers performed by the examinee can receive credit.

The training protocol for SPs is a "building blocks" approach in which trainers meet with a new SP multiple times, each time adding additional case information and details, to avoid overwhelming the SP with too much information at once. The process takes about 1 month, and combines group training (2–10 SPs) and individual training. Each session lasts about 2–3 hours and includes time to review the previous session and materials.

The first step is an initial group orientation session to tell potential SPs who have applied for the job about the exam and what is expected (long hours, multiple physical examinations, and keeping confidential the details of the exam). Each SP has a medical examination and is matched to a case. SPs are told that, if the SP does not successfully complete training, he/she will not be hired for the exam, but will be paid for attending the training. All SPs must sign confidentiality agreements before starting training, which state that they will not discuss the details of the exam with anyone outside of the organization.

Training Session 1 is an individual session, which takes place with a trainer and a new SP. The goal for the first session is for the SP to learn the details of the patient's life and present illness, checklist, and appropriate cues for each checklist item. Using the eCase, the SPs watch videotaped examples of the portrayal and encounters so that the SP can practice completing checklists.

Portrayal requirements are essentially the rules that the SP must follow for standardization. One example of these requirements is that SPs should always stay "in role" when the examinee is in the room. Another portrayal requirement is that the SP cannot give an examinee a checklist item without being prompted or cued by the examinee.

At the completion of the first training session, the SP takes an online quiz about the facts of the case. The trainer can immediately access the results and review the quiz with the SP. These results become part of the SP's permanent record of training in the eCase, which can be accessed by the trainer and center manager.

Training Session 2 builds on the facts learned during Session 1, and expands to include the response to the physical examination and the patient's affect. This session, which takes place with an individual SP and a trainer, begins to bring the elements of the case together.

The trainer takes on the role of "examinee" and conducts multiple role-play exercises with the SP practicing correct responses. Each role play is videotaped, and the tape is reviewed with the SP. The SP is expected to complete checklists for each role-play encounter, with at least 75% accuracy. The accuracy of the checklist is checked by videotape.

Then, another trainer will come into the training to role-play as an examinee for additional practice and, again, with videotaping. Because the SP will get used to the same trainer, bringing in another trainer gives them an idea of how examinees will ask the same questions in different ways.

Communication and Interpersonal Skills (CIS) and Spoken English Proficiency (SEP) Training

Training SPs to complete the CIS and SEP scales is a 3-day process, completed with a group of 2–10 SPs. A videotaped program is used to standardize this training, which is facilitated by a trainer from the site. On day 1, the SPs learn how to complete the CIS scale by watching a video, and are given practice watching videotaped encounters of real examinees, ratings, and the group discussion of the ratings. On day 2, SPs learn how to complete the SEP scale by watching a video, and the SPs practice rating spoken English by watching videotaped encounters. On the final day, the SPs individually rate a collection of 12 videotapes, and scores are checked to ensure the SPs are rating within the acceptable guidelines.

Training Session 3 is similar to Session 2, but with higher expectations for SP performance. SPs are

expected to complete a checklist for each practice encounter in addition to completing the CIS/SEP scales. The SP is expected to be 100% accurate in portrayal and in completing the checklists and rating scales.

The SP "final exam" can be thought of as a dress rehearsal, and is usually done with a group of SPs. The actual setting of the exam session is simulated, so that the SPs can experience the timing of having one examinee after another. The site trainers are assigned various roles to play such as an incompetent examinee, a mechanical examinee, or a rude examinee, all in an effort to mimic actual situations that the SPs will encounter. After the final exam, the videotapes are reviewed by the trainers and SPs. The SPs are then checked for accuracy in their portrayal and completion of the checklists and rating scales, and are given feedback by their trainers.

After completing training, the SP is assigned to an unscored station in the live exam. While in the unscored station, the SP's trainer observes and completes a monitoring checklist, and meets with the SP after each encounter to give them feedback on their portrayal and checklist accuracy.

When the trainer feels the SP is performing and rating according to expectations, they submit a request to the central CSEC office for "sign-off." Sign-off consists of two random encounters by the SP selected for viewing and assessment by a case specialist at the central CSEC offices. The SP must perform up to standards on both encounters to be approved, or "signed-off", to enter the live exam. If not, the SP receives additional training and is resubmitted by the site trainer for sign-off.

Once they are approved to enter the exam, the SP must play in the unscored station until sufficient data are collected by the psychometricians to equate the case as part of scoring. The SP can then be assigned a place in the live exam in a scored station.

The key element vital to training and standardization of SPs is feedback. Feedback is given consistently throughout and after training, through live and video review by the central CSEC offices of the live SP encounters with examinees.

Training the Trainers

SP trainers come from a variety of backgrounds including teachers, former military trainers, former dog trainers, and former SPs. Trainers are responsible for training SPs and for performing quality assurance checks on their SPs, as well as on SPs at other sites.

Before the Step 2 CS examination was introduced, new trainers were trained in groups of four or five; now this training takes place on an individual basis whenever a new trainer is hired. First, the new trainer is assigned a more senior trainer to serve as mentor, and they spend 2–4 weeks at the site, observing and assisting the mentor. In this way, new trainers learn about center operations and observe SP training and quality assurance measures.

Then, the new trainer is sent to the central CSEC office in Philadelphia where they spend 1 week training a new SP to a case in an accelerated way, using the training protocols. A central staff member gives feedback to the new trainer after each step of the training process. Although routine SP training for the exam takes about 1 month, the focus of the trainer's training is not on the performance of the SP but rather on the trainer learning to use the training protocols giving oral and written feedback to the SP about performance.

Quality Assurance

CSEC Quality Assurance (QA) is a large program that encompasses the entire exam from development of case materials to scoring. QA checks are performed on SP performance and rating accuracy, as well as on the scales and checklists in use, the physician note raters, and the examinee scores. For the purpose of this paper, the focus will be on the SP QA portion of the program.

Live review is required by the trainers who watch their SPs while the exam is in progress on a daily basis. Each site trainer and case developer is a case specialist for 5–10 cases. A web-based application assigns videotaped encounters to each case specialist for the SPs playing those cases across sites. This is done so that the case specialist can confirm the performance of the SPs across sites; for example, so that the John Doe case playing at the Los Angeles center looks like the one playing at the Atlanta and Houston center.

While the trainer or case specialist is observing the SP (either live or by video), they complete a monitoring form, which is essentially a copy of the checklist plus portrayal criteria, for example, the correct affect of the case and response to the physical exam.

The trainers complete these forms using a web-based quality assurance tool. The report is received by the central CSEC QA staff, who review all the reports, and send copies to the center managers and trainers for review. The trainer at the site then meets with the SP to review these written reports.

Consistently, over 96% of the SPs perform at a perfect or minor error level. Less than 4% of the SPs required additional training, and 0.1% had to be pulled from the examination. This demonstrates that the training methods are effective in producing standardized, accurate SPs for the Step 2 CS examination.

Summary

Training SP trainers and SPs in a consistent, standardized manner produces accurate portrayal and recording. A "building blocks" process of training SPs using a web-based program can be accomplished in about 1 month. Feedback to the trainers from the central CSEC staff and to the SPs via the trainers is an essential part of the QA program for this exam. The ultimate goal is standardization so that this portion of the US licensure process is reliable and equitable, irrespective of which of the five testing centers administers the exam.

References

1. Barrows H. An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Acad Med* 1993;68:443–51.
2. Colliver JA, Williams RG. Technical issues: test application. *Acad Med* 1993;68:454–63.
3. Rethans JJ, Sturmans F, Drop R, et al. Assessment of the performance of general practitioners by the use of standardized (simulated) patients. *Br J Gen Pract* 1991; 41:97–9.
4. Norman GR, Tugwell P, Feightner JW. A comparison of resident performance on real and simulated patients. *J Med Educ* 1982;57:708–15.
5. Tamblyn RM, Klass DJ, Schnable GK, et al. The accuracy of standardized patient presentations. *Med Ed* 1991;25: 100–9.
6. Furman GE, Brown C, Swygert KA, et al. What traits make a good SP? Results of a personality inventory for SPs working in a high stakes examination. In: *Proceedings of the 13th Ottawa International Conference on Clinical Competence*, Melbourne, Australia, March 8, 2008.