

Using standardised patients in an objective structured clinical examination as a patient safety tool

J B Battles, S L Wilkinson, S J Lee

Qual Saf Health Care 2004;13(Suppl 1):i46–i50. doi: 10.1136/qshc.2004.009803

Standardised patients (SPs) are a powerful form of simulation that has now become commonplace in training and assessment in medical education throughout the world. Standardised patients are individuals, with or without actual disease, who have been trained to portray a medical case in a consistent manner. They are now the gold standard for measuring the competence of physicians and other health professionals, and the quality of their practice. A common way in which SPs are used in performance assessment has been as part of an objective structured clinical examination (OSCE). The use of an SP based OSCE can be a powerful tool in measuring continued competence in human reliability and skill performance where such skills are a critical attribute to maintaining patient safety. This article will describe how an OSCE could be used as a patient safety tool based on cases derived from actual events related to postdonation information in the blood collection process. The OSCE was developed as a competency examination for health history takers. Postdonation information events in the blood collection process account for the majority of errors reported to the US Food and Drug Administration. SP based assessment is an important patient safety tool that could be applied to a variety of patient safety settings and situations, and should be considered an important weapon in the war on medical error and patient harm.

patient safety that links human behaviour within the process of care, nested within the structure of care. The art of measuring the individual performance nested within complex structures and processes of care is an important challenge to patient safety. Both Rasmussen⁴ and Reason⁵ have described different types of human error and system failures that contribute to safety events, and van der Schaaf⁶ and Vincent⁷ have each developed causal classification approaches to apply these concepts as causal classification schema to aid in the analysis of risk and hazards contained in safety events. How are we to detect risks and hazards associated with errors of omission in health professionals where failure to obtain critical patient information can have far reaching consequences?

A great deal of attention in patient safety has been paid to using event reports to identify risks and hazards and potential failures of graduate medical trainees.^{8,9} However, these methods are retrospective and may not be adequate for more prospective competency assessment in areas where human performance may be subject to critical failures increasing risks and hazards to patients.

If we are to shift to a more prospective method of identifying risks and hazards associated with human performance, simulation may be one method to examine human performance without placing actual patients at risk from the very failures we wish to eliminate. One of the most powerful forms of simulation currently in widespread use in medical education is the standardised patient (SP). The use of SPs in both teaching and in evaluation has now become commonplace in most medical schools throughout the world. The SP can be considered a sophisticated form of role playing, a widely used form of simulation.

SPs are individuals, with or without actual disease, who have been trained to portray a medical case or role in a consistent manner.¹⁰ SPs can also evaluate skills in interviewing, interpersonal relationships, and communication.¹¹ The use of SPs is advantageous over real patients in that an SP can be available at any time and setting, and all subjects are exposed to the same situation. SPs have been used in the evaluation and training of medical students, residents, and

Continuing competency assessment after initial licensing and/or certification is an important element in a proactive patient safety environment.¹ An important issue in the area of competency assessment is determining what constitutes valid and reliable methods that focus on actual performance rather than on knowledge assessment only, in the form of written examinations. A recent patient safety report, *Patient Safety: Achieving a New Standard of Care*², stressed the importance of examining not only errors of commission but also those of omission. The issue of identifying risks and hazards related to omissions is particularly challenging, especially in association with individual healthcare professional performance in critical provider–patient interactions, where missed information can have far reaching consequences downstream in the process of care. Battles and Lilford³ have presented a model of

See end of article for authors' affiliations

Correspondence to:
Dr J B Battles, Agency for
Healthcare Research and
Quality, Center for Quality
Improvement and Patient
Safety, 540 Gaither Road,
Rockville, MD 20850,
USA; jbbattles@ahrq.gov

Abbreviations: BPD, blood product deviation; FDA, Food and Drug Administration; G theory, generalisability theory; HXE, history examination extraction; HXI, history examination interpretation; HIV, human immunodeficiency virus; OSCE, objective structured clinical examination; SD, standardised donor; SP, standardised patient

practising physicians. With proper training, SPs have been shown to provide consistent and accurate simulations and recordings of performance by medical students and professionals.¹² SPs are now the gold standard for measuring the competence of physicians and the quality of their practice,¹³ and with adequate training are indistinguishable from real patients.¹⁴

While SP are now widely used in teaching, there has been a parallel growth in SP based performance examinations. One of the most commonly used formats for SP competency based assessment is that of an objective structured clinical examination (OSCE). OSCEs have been used in Canada for certification since the early 1990s. The United States Medical Licensing Examination began using the OSCE format as part of the medical licensing process in the United States in 2004.¹⁵ The OSCE was first described in 1975 by Harden,¹⁶ a pioneer in medical education from Scotland. The OSCE has spread from Scotland to other parts of the UK, North America, Australia, and worldwide. The OSCE is not so much a testing method as it is a flexible approach to test administration. In an OSCE, a variety of methods can be incorporated to obtain an assessment of clinical skills.¹⁶ Examinees in an OSCE rotate around a circuit of stations where they are required to perform a variety of clinical tasks, which could include taking a history from a patient, performing a physical examination, and ordering and interpreting diagnostic studies. Many of these stations incorporate SPs.

Can the lessons learned in using SP based performance assessment be applied to patient safety? We believe that the answer is "Yes"; the SP based OSCE can be an important patient safety tool and we offer the following case example from transfusion medicine to make our point.

COMPETENCY EXAMINATION IN TAKING HEALTH HISTORIES

Concern over the safety of the world's blood supply following the outbreak of the AIDS epidemic has been with us now for nearly 2 decades.¹⁷⁻²¹ In transfusion medicine, post-donation information accounts for the majority of events reported by blood and plasma centres to the US Food and Drug Administration (FDA) through a mandatory reporting system, the blood product deviation (BPD) report.²² The FDA further classifies postdonation information into two categories; information that was known or could have been known at the time of donation, and information not known at the time of the donation.²² Of the 33 466 total BPD reports submitted the FDA in 2002 from blood and plasma establishments, 22 162 (69.2%) were postdonation information reports, 88.7% of which were known by the donor at the time of donation.²³ This information is most commonly discovered at a subsequent donation attempt.

In 1995, the FDA issued the *Guidelines for Quality Assurance in Blood Establishments*.²⁴ These guidelines state that formal competency evaluation programmes analysing both theoretical and practical knowledge of procedures should be implemented. The evaluation could include direct observation, records review, written tests, and internal and external proficiency specimens. We wished to explore a more comprehensive form of competency assessment of our health history takers. Within the USA, donor collection employees undergo an extensive training programme upon hiring. Typically in the USA, at most community blood centres, the donor collection staff is composed of registered nurses, licensed practical nurses, emergency medical technicians, and paramedics. All donor collection staff members are required by FDA regulation to undergo annual competency evaluations.²⁵ While it is often believed that the donor merely failed to provide information that subsequently led to a BPD

report as a result of postdonation information, we wanted to more clearly quantify and understand the role the medical history takers might play regarding these types of reportable events. Implementing the OSCE was an opportunity to examine the predonation history process from a postdonation information focus and to identify improvements that would reduce their occurrence.

APPROACH OSCE design

An eight station OSCE incorporating SPs as standardised donors (SDs) was developed to assess the competency of the health history takers at a typical community based blood centre, Hoxworth Blood Center at the University of Cincinnati Medical Center. The OSCE was adapted from the OSCE of the University of Texas Southwestern Medical Center. It was used to provide performance assessment of medical students in their "Instruction to Medicine" course, given in the second year of a 4 year curriculum.²⁶ The OSCE measured two individual skill components: history taking technique (history examination extraction; HXE), and ability to record and interpret a medical history (history examination interpretation; HXI). The SDs evaluated the HXE component with the use of a checklist. The completed donor forms were used to measure the HXI component. The total number of cases, eight, was set a priori, based on previous experience with OSCE administration from 1989. As Swanson and Norcini²⁷ point out, the greatest threat to reliability in performance examinations is case specificity, where performance on one case does not necessarily predict performance on another, therefore, the number of cases needs to be sufficient to avoid this problem and yet be of a practical length that is reasonable to administer.

Standardised donor scenarios

The eight SD cases for whole blood donation were designed to evaluate the affective and cognitive skills of the person taking the health history. The cases were based on the postdonation accidents seen at Hoxworth Blood Center. Of the 67 incidences of postdonation information Hoxworth Blood Center reported to the FDA from 1 January to 30 June 1999, over half were related to donors who had visited an area endemic for malaria, but not disclosed that information during the predonation interview process. The next highest group included donors who had received a tattoo or body piercing, followed by donors with high risk behaviour for HIV. Because of the recent interest in bloodborne transmission of bovine spongiform encephalopathy and restriction of travel to the UK for donors, a case was developed to reflect this emerging area of potential postdonation information, even though this had not previously been an area of concern.

Table 1 Standardised donor scenarios

No.	Gender	Suitability	Description
1	Female	Acceptable	Hormone replacement therapy. Professional ear piercing. No travel
2	Male or female	Acceptable	Hypertension medication
3	Female	Defer	Allergy medication. Travel to malaria areas in South Africa
4	Female	Defer	Basal skin cancer in 1992. Travel to malaria areas in Belize
5	Female	Defer	Medication for depression. Unprofessional ear piercing
6	Female	Defer	Recent flu shot. Treatment for chlamydia
7	Male	Acceptable	Cholesterol medication. <6 months cumulative travel to the UK
8	Male	Defer	HIV high risk factor

The mix of cases was also matched to the national profile of reported events from the FDA. Table 1 lists the eight cases used.

Standardised donor HXE checklist

A 15 item checklist was developed for the eight cases. The first thirteen of the items were the same for each case, and the last two items were case specific. The items identified specific communication and history taking tasks to evaluate the affective skills (HXE) of the person taking the health history. The items were chosen from literature on medical interviewing and are shown in table 2.²⁸⁻³² While there are differences between the patient–doctor encounter and the health history taker–potential blood donor encounter, many interview techniques apply to both situations. The SDs were trained in assessing and evaluating the performance of the health history takers based on the checklist items. All items were equally weighted. Satisfactory performance on the item was worth one point and somewhat satisfactory performance was worth half a point.

Standardised donors

Twelve SDs were recruited to present the eight cases. Eight of the twelve had previous experience as SPs for the University of Cincinnati Medical School. Each SD received information about blood donation, including a donor form, the case information, and a generic checklist. The SDs were trained to present the essential features and history of each case and to answer probable questions posed by the history takers. The SDs were instructed to answer the questions asked by the person taking their health history but not to volunteer information. They were also trained in evaluating the health history takers and in completing the checklists.

Testing

Over a 1 month period, nine OSCE sessions were conducted at Hoxworth Blood Center, and a total of 56 health history takers were evaluated. At the beginning of each OSCE, instructions were given to each health history taker. The history takers knew that the donors and cases were not “real”. A range of 1–8 health history takers was tested per session. Each cubicle assigned to the individual health history taker contained a copy of Hoxworth Blood Center’s *Donor Selection Standard Operating Procedures*,³³ the high risk questions asked of all donors, the book of *Donor Medical Criteria*,³⁴ and pens. At the beginning of each encounter, the SD gave the history taker a donor form. The SDs that portrayed the three cases involving travel outside the US and Canada carried copies of world atlases and the *Health Information for International Travel, 1999–2000* by the Centers for Disease Control and Prevention,³⁵ and presented them to the screener

at the start of the encounter. Time limits were not placed on the encounters. The health history takers did not take any physical measurements such as blood pressure, pulse, temperature, or haematocrit. The encounter consisted of asking and answering the questions on the donor form, and the history takers were instructed to base their donor suitability determinations on the questions only. At the conclusion of the encounter, the SD left the history taker and completed the checklist (HXE). The process continued until each health history taker had met with all eight SDs.

Donor form grading

The donor forms were evaluated for completeness and for the correct determination of donor suitability. Points were assigned based on the required documentation for each scenario. Each donor form was worth a total of 100 points. The donor forms measured the ability to record and interpret a medical history (HXI).

Statistical analysis

Reliability of the OSCE was determined by calculating generalisability coefficients³⁶ and a dependability index with cut score³⁷ for the entire OSCE and each of the two skill components. The Spearman-Brown prophecy formula was used to determine the number of stations needed to achieve a generalisability coefficient or a dependability of 0.8.³⁸ Statistical analysis was performed using SAS software (version 6.12; SAS Institute, Inc. Cary, NC, USA).³⁹

FINDINGS

A detailed description of the results and psychometric analysis of this OSCE has been reported elsewhere in the literature.⁴⁰ We, will, therefore, only briefly describe the highlights of the results as an example of a patient safety related OSCE.

Health history taker OSCE scores

Because all 56 history takers participating in the OSCE are employed health professionals who determine donor suitability on a routine basis, high scores for the HXI component were anticipated and expected. A case could be made for the HXI score to be 100 in order to demonstrate competency; however, the cases were designed to test the competency level of all staff. As most of the cases were derived from previous postdonation accident information, determining donor suitability was already shown to be difficult and not a perfect process. While a 100% performance on HXI is highly desirable, and while 25 out of 56 subjects achieved it, the test might have appeared too easy had all the subjects scored 100%.

For most of the history takers, the HXI scores were greater than the HXE scores. In other words, subjects performed better on completing the donor form than on interviewing skills. Ten subjects had an HXE score of <90.00; four of these had HXI scores of 100. The largest range of scores for a single subject was an HXI of 100 with HXE of 80.8. These subjects correctly determined donor suitability but obviously some improvement is needed in their interview and interpersonal skills.

Only one person received an HXE score of 100, and no one received a score of 100 on both components. The history taker with the lowest score for HXE also had the lowest score for HXI. In some of the case scenarios, the history takers obtained the correct answer but did not endeavor themselves to the SDs. One SD commented that one of the history takers made no attempt at personal warmth. How likely would a real donor be to return in this instance? For the subject with the lowest scores in both components, their poor interviewing skills prohibited them from receiving the necessary information to correctly determine donor suitability. Lack of

Table 2 Scores for checklists: common items

Checklist items		Score (max = 448)	Percentage
No.	Item		
1	Introduced themselves	317.5	70.9
2	Warm and friendly	430.5	96.1
3	Poised and confident	426.5	95.2
4	Vocal qualities	417.5	93.2
5	Nonverbal communication	425.5	95.0
6	Listening skills	426.5	95.2
7	Maintain donor comfort	425.5	95.0
8	Interview style	432.0	96.4
9	Answers recorded immediately	436.0	97.3
10	Clarifying questions	430.0	96.0
11	Confident with decision	426.0	95.2
12	Explained eligibility/deferral	433.0	96.7
13	Appreciation	430.0	96.0

knowledge or limited facility in a language (mismatch between patient' and provider's primary language) can be a significant issue in the execution of interviewing skills.

Generalisability and dependability

The reliability of this OSCE as determined by generalisability coefficient did not reach the desired level of 0.8. This is common for standardised performance based assessments;^{36, 37} the test would have to double in size in order to attain the desired level. The test length required for reliable results is a well documented drawback of OSCEs.³⁶

The reliability increased when dependability index with cut score was used. This was anticipated and has been reported in the literature.⁴¹ The reliability was >0.8 for the total OSCE score and the HXI component with this method. Using a cut score of 0.9, or 90%, 10 subjects would not have passed the exam based on the HXE score, and one person would not have demonstrated competency based on the HXI score.

Competency determination and retraining

Results of the OSCE were reported to the donor collection management team and staff. The failure to defer the SD who had visited South Africa (case 3) and the SD with chlamydia (case 6) were viewed as the most severe errors in the OSCE. From a recipient safety perspective, failing to defer an unsuitable donor is more serious than deferring an acceptable donor. Three history takers were involved in the four incidences and they were retrained by the donor collection department supervisor and re-evaluated. Re-evaluation for the malaria deferral consisted of a written test that involved obtaining information from the *Health Information for International Travel, 1999–2000*.³⁵ The history taker who also missed the chlamydia deferral had to take an additional written test that required completing donor forms for three written scenarios. Two of the three history takers had the two lowest total scores for the OSCE. The third had an overall OSCE score of 93.452. The history taker with the lowest score is not from the USA and all eight SDs commented that he had language problems, and was difficult to understand. Because he did not correctly determine donor suitability owing to poor history taking, the donor must now read the questions from the form along with him when he screens potential blood donors.

While correctly determining donor suitability is the main goal of the pre-donation screening, it is not the only determination of a successful interview. Several health history takers correctly completed the donor forms but had low communication scores. In a real donation setting, even if the donor was correctly deferred or accepted, one negative interview experience might result in the donor failing to return to donate blood.

Comparing results with other OSCEs

The overall template for the design and administration of this examination demonstrating competency in taking donor history was based on the established OSCE literature and many years of experience in conducting OSCEs in the medical education arena. Thus, we were able to compare the results of this OSCE to others that the primary author had administered. The results of this OSCE matched the patterns results and psychometric results that had previously been achieved.^{26, 42, 43} Achieving this level of comparability was gratifying and encouraging, indicating that the attributes of SP based competency examinations are domain independent.

LESSONS

Given the number of BPD reports stemming from postdonation information that are reported to the FDA, the competency of health history takers is an important issue that must be addressed in any discussion in maintaining the safety of

the blood supply. The OSCE format represents a valid and reliable format for conducting annual competency assessment as required by the FDA.

There are a number of other important lessons that can be learned from this example of a transfusion based OSCE used as a patient safety tool. Firstly, the OSCE works well at measuring clinical performance in a critical area for patient safety. Secondly, this transfusion example is a confirmation that the standards and protocols for design and data analysis of SP based OSCEs that have emerged from medical education work well in other domains, and for health professionals other than physicians in training. The psychometric protocols already exist for determining the reliability and test performance of such examinations. Thirdly, available resources that can be used to assist in the development and employment of this type of simulation based assessment examinations exist in many settings. Many medical schools have a cadre of individuals who help recruit and train SPs for both teaching and assessment. Those in patient safety are not alone; there is help if providers know where to look for it, especially in medical education units.

The OSCE is a template for a competency assessment examination. While much of the work to date has featured OSCE stations that employ SP and human simulators, this does not mean that nonhuman simulators cannot be used in this format. There are important limitations in using SPs. Assessing invasive procedures is not something that is appropriate or ethical for SPs. In such cases, models and other mechanical or electronic simulators would be the choice. However, the assessment templates and psychometric standards that have emerged from the use of SP based OSCE can be applied to performance assessments with other types of simulators. Thus, advocates of high fidelity simulators should look to the OSCE evidence based literature for help and guidance in designing and evaluating simulator based competency assessment.

There are a number of other applications for SP to be employed in the process of identifying risks and hazards in patient safety. As Norman¹² and colleagues pointed out almost 20 years ago, SPs are indistinguishable from actual patients in actual practice settings, so could be used to detect errors of omission in ambulatory practice settings. SPs may also be able to be used to examine transitions of care where

Key messages

- SP based OSCEs are valid and reliable tools that can be applied to a variety of patient safety settings to assess continuing competency.
- Extensive literature based evidence to support the design, implementation, and psychometric analysis of results of OSCEs exists, for both SPs and other forms of simulation.
- SPs can be used in other ways to assess the risks and hazards associated with error of omission in the process of care, especially in ambulatory care settings.
- Resources exist to support use of SPs and OSCEs. Many medical schools throughout the world have developed resources to train and recruit SP and conduct OSCEs. These resources could help those in patient safety adopt these assessment techniques
- SPs and OSCE are no longer only for medical education. These forms of simulation can and are being used for licensure, recertification, and continuing competency assessment in a variety of areas.

an SP would move from one setting to another. These are just two examples of how SPs can be an effective form of simulation that could be part of a patient safety programme.

Many of the methods and approaches to patient safety have come from other industries; in the case of the SP based OSCE, the methods, resources, and expertise already exist within healthcare – we just need to make use of the available resources. The OSCE and SPs are no longer only for medical education.

Authors' affiliations

J B Battles, United States Department of Health and Human Services, Agency for Healthcare Quality and Research, Center for Quality Improvement and Patient Safety, Rockville, MD, USA

S L Wilkinson, Hoxworth Blood Center, University of Cincinnati Medical Center, Cincinnati, OH, USA

S J Lee, The Blood Center of Southeastern Wisconsin, WI, USA

Partial funding provided by Grant # R01HL53772 from the United States National Heart, Lung and Blood Institute, National Institutes of Health.

Disclaimer: The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Agency for Healthcare Research and Quality, the United States, Department of Health and Human Services.

At the time this research was conducted, Dr. Battles was Professor of Medical Education at the University of Texas Southwestern Medical Center at Dallas and co-principal investigator for the Medical Event Reporting System for Transfusion Medicine (MERS-TM) project.

REFERENCES

- In: **Kohn LT**, Corrigan JM, Donaldson MS, eds. *To err is human: building a safer health system*, Washington, D.C.: National Academy Press, 2000.
- In: **Aspden P**, Corrigan JM, Wolcott J, Erickson SM, eds. *Patient safety: achieving a new standard of care*. Washington, DC: National Academy Press, 2003.
- Battles JB**, Lilford RJ. Organizing patient safety research to identify risks and hazards. *Qual Saf Health Care* 2003;**12**(suppl 1):ii2-7.
- Rasmussen J**. Outlines of a hybrid model of the process operator. In: Sheridan TB, Johansen G, eds. *Monitoring behavior and supervisory control*. New York: Plenum Press, 1976.
- Reason J**. *Human error*. Cambridge, GB: Cambridge University Press, 1990.
- Van der Schaaf TW**. *Near miss reporting in the chemical process industry*. PhD Thesis, Eindhoven, NL: Eindhoven University of Technology, 1992.
- Vincent C**, Taylor-Admas S, Chapman EJ, et al. How to investigate and analyze clinical incidents: clinical risk unit and association of litigation and risk management protocol. *BMJ* 200, **320**:777-81.
- Battles JB**, Kaplan HS, Van der Schaaf TW, et al. The attributes of medical event-reporting systems: experience with a prototype medical eventreporting system for transfusion medicine. *Arch Pathol Lab Med*, 1998 Mar, **122**:231-8.
- Battles JB**, Shea CE. A system of analyzing medical errors to improve GME curricula and programs. *Acad Med* 2001;**76**:125-33.
- Stillman PL**. Session three—technical issues: logistics. *Acad Med* 1993;**68**:464-70.
- Vu NV**, Barrows HS. Use of standardized patients in clinical assessments: recent developments and measurement findings. *Educ Researcher* 1994;**23**:23-30.
- Norman GR**, Neufeld VR, Walsh A, et al. Measuring physicians' performances by using simulated patients. *J Med Educ* 1985;**60**:925-34.
- Peabody JW**, Luck J, Glassman P, et al. Comparison of vignettes, standardized patients and chart abstraction. *JAMA* 2000;**283**:1715-22.
- Researchers of Clinical Skills Assessment**. Consensus statement of the Researchers of Clinical Skills Assessment (RSCA) on the use of standardized patients to evaluate clinical skills. *Acad Med* 1993;**68**:475-7.
- National Board of Medical Examiners (NBME)**. 2004 USMLE step 2 clinical skills update. Available at: www.usmle.org/step2/step2cs/Step2CS2004update.asp.
- Harden R**, Stevenson M, Downie W, et al. Assessment of clinical competence using objective structured examinations. *Br Med J* 1975;**1**:447-51.
- HTLV-III Consensus Conference**. Consensus conference: the impact of routine HTLV-III antibody testing of blood and plasma donors on public health. *JAMA* 1986;**256**:1178.
- Marconi M**, Almini D, Pizzi MN, et al. Quality assurance of clinical transfusion practice by implementation of the privilege of blood perception and computerized prospect vs audits of blood requests. *Transfusion* 1996;**6**:11-19.
- McCelland DBL**, McMenamin JL, Moores HM, et al. Reducing risks in blood transfusion: process and outcome. *Transfusion* 1996;**6**:1-10.
- Gambino R**, Mallon P. Near misses—an untapped data base to find root causes. *Lab Report* 1991;**13**:41-4.
- Murphy WG**, McClelland DBL. Deceptively low morbidity from failure to practice safe blood transfusion: an analysis of serious blood transfusion errors. *Vox Sang* 1989;**57**:59-62.
- Food and Drug Administration**. *Biological product deviation report, Form FDA 3486*. March, 2001.
- Food and Drug Administration, Center for Biologics Evaluation and Research**. *Error and accident reports—summary for FY2002*, Memorandum, March, 2004.
- Food and Drug Administration**. *21 CFR Parts 600 and 606 Final Rule*, 7 November, 2001.
- Food and Drug Administration, Center for Biologics Evaluation and Research**. *Guideline for quality assurance in blood establishments*, 11 July, 1995.
- Carpenter JL**, McIntire D, Battles J, et al. Administration of a parallel, simultaneous objective structured clinical examination to accommodate a large class of students. *Teach Learn Med* 1993;**5**:79-85.
- Swanson DB**, Norcini JJ. Factors influencing reproducibility of tests using standardized patients. *Teach Learn Med* 1989;**1**:158-66.
- Billings JA**, Stoekle JD. *The clinical encounter: a guide to the medical interview and case presentation*, 2nd ed. St. Louis: Mosby, 1999.
- Lazare A**, Putnam SM, Lipkin M. Three functions of the medical interview. In: Lipkin M, Putnam SM, Lazare A, eds. *The medical interview: critical care, education, and research*. New York: Springer-Verlag, 1995:3-19.
- Aldrich CK**. *The medical interview: gateway to the doctor-patient relationship*, 2nd ed. New York: Parthenon, 1999.
- Lipkin M**, Frankel RM, Beckman HB, et al. Performing the interview. In: Lipkin M, Putnam SM, Lazare A, eds. *The medical interview: critical care, education, and research*. New York: Springer-Verlag, 1995:65-82.
- Coulehan JL**, Block MR. *The medical interview: mastering skills for clinical practice*, 3rd ed. Philadelphia: FA Davis Company, 1997.
- Hoxworth Blood Center**. *Standard operating procedure B2.000, donor selection*. Cincinnati: University of Cincinnati Press, 2000.
- Hoxworth Blood Center**. *Donor medical criteria, reference list and medication deferral/acceptability criteria list*. Cincinnati: University of Cincinnati Press, 1998.
- Centers for Disease Control and Prevention**. *Health information for international travel, 1999-2000*, US Department of Health and Human Services, 1999.
- Van der Vleuten CPM**, Swanson DB. Reliability issues in objective structured clinical examinations. In: Hart IR, Harden RM, Des Marchais J, eds. *Current developments in assessing clinical competence*. Montreal: Can-Heal, 1992:66-78.
- Shavelson RJ**, Webb NM. *Generalizability Theory A Primer*. Newbury Park: Sage Publications, 1991.
- Brennan RL**, Kane MT. An index of dependability for mastery tests. *J Med Educ* 1977;**14**:277-89.
- SAS Institute**. *SAS/STAT v6.12, 1999*, Cary, SAS Institute Inc., 1999.
- Lee SJ**, Wilkinson SL, Battles JB, et al. An objective structured clinical examination to evaluate health historian competencies. *Transfusion* 2003;**43**:34-41.
- Colliver JA**, Williams RG. Technical ISSUES: TEST APPLICATIONS. *Acad Med* 1993;**68**:454-63.
- Battles JB**, Carpenter JL, McIntire DD, et al. Analyzing and adjusting for variables in a large-scale standardized-patient examination. *Acad Med* 1994;**69**:370-6.
- Colliver JA**, Swartz MH. Assessing clinical performance with standardized patients. *JAMA* 1997;**278**:790-1.