

How to assess communication, professionalism, collaboration and the other intrinsic CanMEDS roles in orthopedic residents: use of an objective structured clinical examination (OSCE)

Tim Dwyer, MBBS*†
 Susan Glover Takahashi, MA, PhD‡
 Melissa Kennedy Hynes, MA‡
 Jodi Herold, PhD§
 David Wasserstein, MD§
 Markku Nousiainen, MD, MSc,
 MEd¶
 Peter Ferguson, MD†
 Veronica Wadey, MD, MA**
 M. Lucas Murnaghan, MD,
 MEd***
 Tim Leroux, MD§
 John Semple, MD, MSc*
 Brian Hodges, MD, Med, PhD§§
 Darrell Ogilvie-Harris, MD,
 MSc****¶¶

From the *Women's College Hospital, †Mt. Sinai Hospital, ‡Postgraduate Medicine Office, University of Toronto, §University of Toronto, ¶Holland Orthopaedic and Arthritic Centre, **Sunnybrook Health Sciences Centre, ††University of Toronto Orthopaedics Sports Medicine, ‡‡The Hospital for Sick Children, §§Toronto General Hospital, and ¶¶Toronto Western Hospital, Toronto, Ont.

Presented at the International Conference Surgical Education and Training (ICOSET) 2013, Canadian Orthopaedic Association 2013, Canadian Conference Medical Education 2013.

Accepted for publication
 Oct. 29, 2013

Correspondence to:
 T. Dwyer
 Women's College Hospital
 76 Grenville St.
 Toronto ON M5S 1B1
 dwyer.orthopedics@gmail.com

DOI: 10.1503/cjs.018813

Background: Assessing residents' understanding and application of the 6 intrinsic CanMEDS roles (communicator, professional, manager, collaborator, health advocate, scholar) is challenging for postgraduate medical educators. We hypothesized that an objective structured clinical examination (OSCE) designed to assess multiple intrinsic CanMEDS roles would be sufficiently reliable and valid.

Methods: The OSCE comprised 6 10-minute stations, each testing 2 intrinsic roles using case-based scenarios (with or without the use of standardized patients). Residents were evaluated using 5-point scales and an overall performance rating at each station. Concurrent validity was sought by correlation with in-training evaluation reports (ITERS) from the last 12 months and an ordinal ranking created by program directors (PDs).

Results: Twenty-five residents from postgraduate years (PGY) 0, 3 and 5 participated. The interstation reliability for total test scores (percent) was 0.87, while reliability for each of the communicator, collaborator, manager and professional roles was greater than 0.8. Total test scores, individual station scores and individual CanMEDS role scores all showed a significant effect by PGY level. Analysis of the PD rankings of intrinsic roles demonstrated a high correlation with the OSCE role scores. A correlation was seen between ITER and OSCE for the communicator role, while the ITER medical expert and total scores highly correlated with the communicator, manager and professional OSCE scores.

Conclusion: An OSCE designed to assess the intrinsic CanMEDS roles was sufficiently valid and reliable for regular use in an orthopedic residency program.

Contexte : Évaluer la compréhension et l'application des 6 rôles intrinsèques CanMEDS (communicateur, professionnel, gestionnaire, collaborateur, promoteur de la santé, érudit) chez les résidents pose un défi pour les responsables de la formation médicale postdoctorale. Nous avons émis l'hypothèse selon laquelle un examen clinique objectif structuré (ECOS) conçu pour évaluer plusieurs rôles CanMEDS intrinsèques serait suffisamment fiable et valide.

Méthodes : L'ECOS comportait 6 stations de 10 minutes, permettant chacune d'évaluer 2 rôles intrinsèques à l'aide de scénarios basés sur des cas (avec ou sans recours à des patients standardisés). Les résidents ont été notés au moyen d'échelles en 5 points et d'une évaluation globale de leur rendement à chacune des stations. La validité convergente a été vérifiée par corrélation avec les rapports d'évaluation en cours de formation (RÉF) des 12 mois précédents et un classement chiffré créé par les directeurs du programme (DP).

Résultats: Vingt-cinq résidents des années 0, 3 et 5 y ont participé. La fiabilité interstation pour les scores totaux aux tests (en pourcentage) a été de 0,87, tandis que la fiabilité pour chacun des rôles de communicateur, collaborateur, gestionnaire et professionnel, a été supérieure à 0,8. Les scores totaux aux tests, les scores aux stations individuelles et les scores pour les rôles CanMEDS individuels ont tous fait état d'un effet significatif selon le niveau des résidents. L'analyse des classements établis par les DP quant aux rôles intrinsèques a révélé une forte corrélation avec les scores au test ECOS pour les rôles. On a observé une corrélation entre les RÉF et l'ECOS pour le rôle de communicateur, tandis que les RÉF pour le rôle d'expert médical et les scores totaux ont été en forte corrélation avec les scores de l'ECOS pour les rôles de communicateur, de gestionnaire et de professionnel.

Conclusion : Un ECOS conçu pour évaluer les rôles CanMEDS intrinsèques s'est révélé suffisamment valide et fiable pour un usage régulier dans un programme de résidence en orthopédie.

The 7 CanMEDS competencies (medical expert, communicator, collaborator, manager, health advocate, scholar and professional) have been clearly outlined in the CanMEDS 2005 Physician Competency Framework, by the Royal College of Physicians and Surgeons of Canada.¹ A similar framework has been described by the Accreditation Council for Graduate Medical Education (ACGME), defining 6 core competencies.² Each of these frameworks describe the principal generic abilities of physicians in health care, and are an integral component of postgraduate education. However, despite the widespread popularity of CanMEDS and other competency frameworks and the mandate to both teach and assess these competencies, the best methods of teaching them remain unknown.^{3,4}

Assessment options for the intrinsic roles (those other than the medical expert role) include in-training evaluation reports (ITERs), structured oral examinations, 360° assessments and objective structured clinical examinations (OSCE).⁴ A survey of a wide variety of medical and surgical program directors in Canada identified that the ITER is the most commonly used method to evaluate the CanMEDS roles, despite its acknowledged subjective nature.⁵⁻⁷ Respondents reported dissatisfaction with current methods of evaluating the intrinsic roles, especially the manager and health advocate roles.

The “OSCE” is a term used to describe a variety of multi-station examinations and is a format currently favoured at orthopedic certification examinations in Canada and other countries. Studies using OSCEs to assess the role of medical expert have demonstrated reliability and validity in postgraduate physician training,⁸⁻¹¹ with some literature demonstrating the ability of the OSCE to assess communication skills.¹²⁻¹⁷ In fact, there is evidence that improved interpersonal skills are linked to better overall OSCE performance.¹⁸⁻²⁰

There is also some evidence that an OSCE can be used to assess other CanMEDS roles, including the scholar role (application of evidence-based medicine or demonstration of teaching skills)^{21,22} and the professional role (cultural awareness and the application of ethical principles).^{23,24} An OSCE has also been used to assess multiple CanMEDS competencies in other fields of postgraduate training, such as radiology and neonatology.^{6,25}

To our knowledge, no research exists regarding methods of assessing the intrinsic roles in orthopedic postgraduate training. We hypothesized that an OSCE designed to assess multiple intrinsic CanMEDS roles would have sufficient reliability and validity to distinguish between different years of postgraduate training in orthopedic residents.

METHODS

Exam development

The orthopaedic residency program at the University of Toronto, in collaboration with the Postgraduate Medical Education Department, designed an orthopedic OSCE to

test the 6 intrinsic CanMEDS roles. A focus group of academic orthopedic specialists was assembled with the goal of creating clinical scenarios evaluating selected CanMEDS principles for each of the roles. The focus group relied on the source document from The Royal College of Physicians and Surgeons of Canada, in which each of the roles, and their key competencies, is clearly defined.¹

The OSCE was 1 hour long and comprised 6 10-minute stations. The roles of communicator, collaborator, professional, manager, health advocate and scholar were assessed; a deliberate attempt was made to avoid testing the medical expert role. Most of the 6 case-based scenarios were designed to assess a primary and secondary role. Stations 2-4 used standardized patients (SPs; station 2: relative concern regarding delay in surgery; station 3: grandmother of child with suspected nonaccidental injury; station 4: teenager being informed of osteosarcoma diagnosis), whereas station 5 used a standardized health professional (SHP; operating room manager). Two stations did not have an SP or SHP (station 1: ethical approach to needlestick injuries; station 6: evidence-based medicine in spinal surgery). Table 1 lists the roles and associated key competencies tested in each station.

The CanMEDS OSCE development was facilitated by an exam blueprint and case development guides. A member of the focus group was assigned as the lead to design each station, which was then reviewed by the entire focus group. Any discrepancies or ambiguities were addressed or removed. A number of 5-point performance rating scales were developed for each of the intrinsic roles. The ratings were anchored by descriptions of performance to be demonstrated by the residents for each role. An overall 5-point global rating was also assigned for each resident at the end of the station. The SPs and SHPs were selected from an established standardized patient bank at the University of Toronto. For the OSCE, 2 SPs/SHPs were trained for each of the stations by an experienced SP trainer; each received a minimum of 3 hours' training for each role. No SP or SHP assessment of performance was used.

Study design

Convenience sampling was used to recruit residents from specific postgraduate years (PGY) of training: PGY0 (incoming residents who had not yet begun training), PGY3 residents and PGY5 residents (volunteers who asked to sit the OSCE). The PGY5 residents had all recently passed their orthopedic certification examinations, and were included as the “gold standard.”

Members of the orthopedic faculty at each station evaluated residents independently. It was not possible to blind examiners from the year of training of the residents, as many of the residents were familiar to the staff surgeons. However, examiners were asked to disregard the year of training when making assessments. The OSCE was

Table 1. Breakdown of each station by primary and secondary roles and key competencies tested

Station; role	Key competencies tested*
Needlestick	
Professional	<ul style="list-style-type: none"> • Bioethical principles • Informed consent/confidentiality • Commitment to professional standards
Manager	<ul style="list-style-type: none"> • Priority setting • Time management
Trauma list	
Manager	<ul style="list-style-type: none"> • Task prioritization and time management • Leadership • Appropriate use of resources in hospital
Communicator	<ul style="list-style-type: none"> • Effective listening • Empathy • Patient-centred approach to communication
Nonaccidental injury	
Health advocate	<ul style="list-style-type: none"> • Recognition of patient risk factors • Modification of risk factors • Patient safety
Manager	<ul style="list-style-type: none"> • Leadership • Negotiation
Breaking bad news	
Communicator	<ul style="list-style-type: none"> • Breaking bad news • Addressing end of life issues • Empathy
Professional	<ul style="list-style-type: none"> • Autonomy in decision making • Bioethical principles
Interacting with OR team	
Collaborator	<ul style="list-style-type: none"> • Conflict resolution • Respect for members of the health care team • Recognizing one's own roles and limitations
Spinal evidence	
Scholar	<ul style="list-style-type: none"> • Evidence-based medicine • Critical appraisal of evidence • Translating knowledge into practice
Communicator	<ul style="list-style-type: none"> • Interactive process • Efficiency and accuracy

OR = operating room.
*Based on the CanMEDS 2005 Physician Competency Framework.¹

Table 2. Resident distribution, by PGY level

PGY	No. of residents
PGY0	6
PGY3	13
PGY5	6
Total	25

PGY = postgraduate year.

Table 3. Total test scores and overall rating of performance

Test score	Group; mean ± SD			
	Grand	PGY0	PGY3	PGY5
Total, out of 100	75% ± 12.9	62.4% ± 5	73.4% ± 9	91.1% ± 8.3
Overall rating of performance, out of 5	3.55 ± 0.78	2.75 ± 0.35	3.45 ± 0.50	4.56 ± 0.44

PGY = postgraduate year; SD = standard deviation.

conducted in 4 1-hour sessions over the course of a single day. Each candidate signed a consent form permitting the use of exam results for research purposes. On completion of the exam, residents were invited to provide feedback using a 5-point Likert scale. Summative and formative feedback was given to each resident at the end of the OSCE.

Concurrent validity was sought in 2 ways. First, we obtained the ITERs from the preceding 12 months for the PGY3 and PGY5 residents, and the results on the 6 intrinsic roles correlated with the OSCE total score and role scores. Second, the 2 program directors (PDs) formed an ordinal ranking of the residents in PGY3 and PGY5 and rated each resident's ability in each of the CanMEDS roles using a 5-point scale (1 = needs significant improvement; 2 = below expectations; 3 = solid, competent performance; 4 = exceeds expectations; 5 = superb). The overall ranking and the rating for each role were also correlated to the total OSCE score and role scores.

Approval for this study was obtained from the Research Ethics Board, University of Toronto. Each resident signed a consent form to permit the use of the OSCE results and ITER results for research purposes.

Statistical analysis

All data were deidentified, and residents were assigned a study-specific number. Raw scores from the individual station scores and role scores were entered into a spreadsheet and analyzed using SPSS version 19. All scores were converted into a percentage, with results reported as mean ± standard deviation. Reliability was established using the interstation α coefficient of reliability (Cronbach α) for each of the scoring tools. We evaluated scores from the different rating scales using regression analysis. The effect of PGY on total test scores (%), overall ratings of performance, individual station scores and role scores were evaluated using 1-way analysis of variance (ANOVA). We considered results to be significant at $p < 0.05$. We used the Scheffe test for post hoc analysis to understand differences in scores between each possible pair of years of training. We determined the correlation between total OSCE scores and role scores with ITER role scores and PD rankings using Pearson correlation and Spearman Rho (R^2), and the Student t test was used to compare the PD ratings of resident performance between PGY levels.

RESULTS

Twenty-five residents from PGY0, 3 and 5 took part in the OSCE (Table 2). The roles of communicator, manager and professional were assessed in multiple stations; collaborator, health advocate and scholar were assessed in 1 station each. The total test scores (converted to a percentage) and the mean overall rating of performance are shown in Table 3; ANOVA testing demonstrated a significant

difference of the effect of PGY on both scores ($p < 0.001$). We found a significant difference between PGY0 and PGY3 ($p = 0.039$), PGY3 and PGY5 ($p = 0.001$), and PGY0 and PGY5 ($p < 0.001$).

The interstation reliability (percent) was 0.87 for total test scores and 0.83 for overall ratings of performance. The internal consistency for 4 of the 6 role scores is shown in Table 4; the consistency for each of these 4 roles was very high (> 0.80). We were not able to compute internal consistency coefficients for the scholar and advocate roles, as only 1 rating scale was used for each of these roles.

The total test scores for the individual stations by PGY are displayed in Figure 1. The effect of PGY on the individual station scores was significant (stations 1, 5 and 6, all $p < 0.01$; stations 2 and 4, both $p < 0.05$) with the exception of station 3 ($p = 0.07$). Post hoc analysis demonstrated a significant difference in station scores between the PGY5 and PGY0 groups and between the PGY5 and PGY3 groups for all stations except station 3. No significant difference in scores was seen between the PGY0 and PGY3 groups, but there was a trend for increased scores in the PGY3 group.

The total test scores for each of the intrinsic roles by PGY are shown in Figure 2. The ANOVA testing for the effect of PGY on each of the role scores was significant (communicator, collaborator, manager and professional roles, all $p < 0.001$; health advocate and scholar roles, both $p < 0.05$). For each of the role scores, the PGY0 and PGY3 groups differed significantly from the PGY5 group, but not from each other.

Analysis of the PD ratings of intrinsic roles demonstrated a good correlation between these and the corresponding OSCE role scores (Table 5). The ITERs from 12 months before the OSCE were available for the PGY3 and PGY5 residents. There was no correlation between ITERs and OSCE scores within any role except for communicator (0.64); however, the ITER overall scores correlated with the communicator (0.58), manager (0.51) and professional (0.56) OSCE role scores.

There was a 64% (16 of 25) response rate to the resident survey. Overall, 87.6% of respondents agreed or strongly agreed that the scenarios reflected encounters that an orthopedic surgeon would have to deal with in general practice, and 81.3% agreed or strongly agreed that participating in the OSCE would help prepare them for their final Royal College examination. However, only 56.3% agreed or strongly agreed that the OSCE was an effective way to assess their understanding of each of the CanMEDS roles.

Table 4. Internal consistency for the 4 role scores with more than 1 rating scale

Role	α Coefficient	Item numbers
Communicator	0.91	17 items in 5 of 6 stations
Collaborator	0.96	3 items in 1 of 6 stations
Manager	0.83	5 items in 3 of 6 stations
Professional	0.84	3 items in 2 of 6 stations

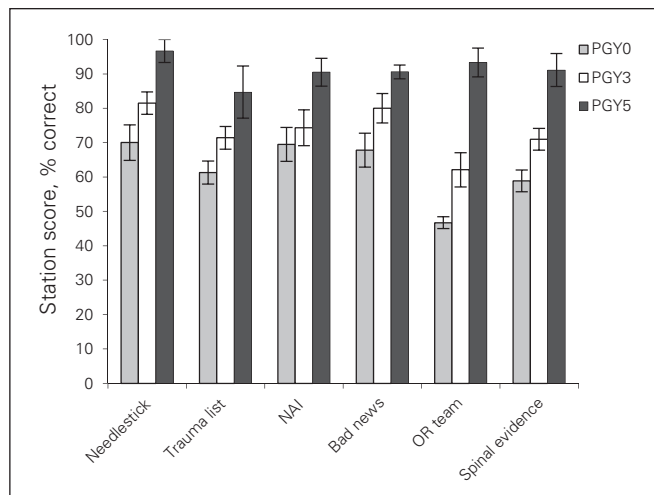


Fig. 1. Total station scores (% correct) by postgraduate year (PGY) for each of the stations. Each station showed a significant difference by PGY ($p < 0.05$) except for station 3 (nonaccidental injury [NAI], $p = 0.07$). Error bars represent standard error of the mean.

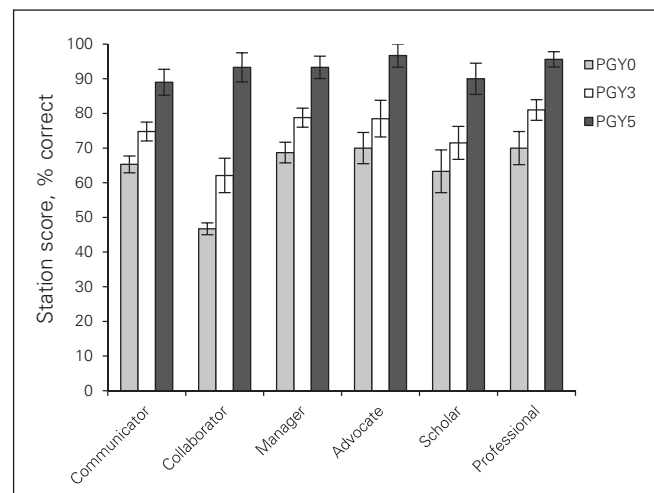


Fig. 2. Total test scores (% correct) for each of the CanMEDS roles by postgraduate year (PGY). Analysis of variance showed significant differences for all roles (communicator, collaborator, manager, professional, all $p < 0.001$; advocate, scholar, both $p < 0.05$).

Table 5. Correlation between the program directors' ratings of resident ability in each of the intrinsic roles and the corresponding OSCE role score

PD role rating	OSCE role score
Communicator	0.79
Collaborator	0.65
Manager	0.66
Health advocate	0.74
Scholar	0.70
Professional	0.72

OSCE = objective structured clinical examination; PD = program director.

DISCUSSION

This orthopedic OSCE, designed to test the 6 intrinsic CanMEDS roles, has shown excellent overall reliability, as well as excellent reliability for the communicator, collaborator, professional and manager roles. Furthermore, using role-specific global ratings, we demonstrated an ability to distinguish between orthopedic residents at different levels of training. To our knowledge, this is the first time formal assessment of intrinsic roles has been studied in the field of postgraduate orthopedic surgical training.

With regards to total test scores, each of the stations were able to demonstrate a statistically significant difference by year of training, with the exception of station 3. Station 3 was a case-based scenario using a standardized patient — a grandmother who has brought in a child thought to have sustained a nonaccidental injury. Residents were asked to take a focused history regarding the home situation and explain to the grandmother the need to alert the appropriate authorities and admit the child. Despite the fact that the PGY effect on station 3 scores was not statistically significant ($p = 0.07$), the scores demonstrate the same general trend as all other stations: PGY5 (mean 90.5%) > PGY3 (mean 74.3%) > PGY1 (mean 69.5%). It may be that this station did not achieve significance because the PGY0 and PGY3 residents performed well, suggesting that these competencies may have been covered in undergraduate medical programs.

Careful blueprinting was used in this CanMEDS OSCE to avoid redundancy. Roles were spread among stations, and the stations that assessed the same roles focused on different competencies within that role, as outlined in the CanMEDS 2005 Framework.¹ For example, in this OSCE, 2 stations (needlestick and trauma list) both examined the roles of communicator and manager, with the needlestick station additionally examining resident understanding of the professional role (bioethical principles and informed consent). However, the trauma list station focused on the competencies of priority setting and time management within the manager role, while the needlestick station sought to examine the competency of managing practice and career effectively.

The OSCE has been previously shown to be a valid and reliable tool for the assessment of the medical expert role, with some evidence for its use in assessing the communicator role.¹²⁻¹⁴ Improved communication skills have previously been linked to both advanced year of training¹⁸ and to increased clinical competence.¹⁹ The OSCE has also been adapted to assess competencies within the roles of professional^{23,24,26} and scholar,^{21,22} with varying amounts of success. For example, Singer and colleagues,²⁴ in an OSCE designed to assess clinical ethics, found a low reliability with only 4 stations; it was felt that increasing the number of stations would be required to obtain acceptable reliability.

Jefferies and colleagues⁶ recently demonstrated that the OSCE may be a valid and reliable method of simultaneously assessing multiple competencies in neonatal/perinatal medicine. Subspecialty trainees were assessed using a combination of binary checklists, 5-point CanMEDS ratings, as well as SPs' and SHPs' assessments of interpersonal and communication skills. Interstation reliability was acceptable to excellent for 6 of the 7 roles, with the scholar role being the exception. Only the teaching component of the scholar role was assessed; the authors recommended creating a single station to assess the competencies inherent to the scholar role, including the ability to understand and evaluate research. We applied this technique with success; our scholar station, designed to assess application of evidence-based medicine, was able to distinguish between residents with different levels of training.

Jefferies and colleagues³ also studied the use of the structured oral examination in assessing the 7 CanMEDS roles, including the medical expert role. Interstation reliability was acceptable for the roles of medical expert, scholar and professional (0.6–0.8), but not for the roles of communicator, collaborator or health advocate (0.4–0.6) or for the role of manager (0.19). In comparison to their previous OSCE study, interstation reliability was lower for all roles except scholar. However, costs were reduced significantly by not using standardized patients. Although we felt that SPs were an important component of our OSCE, the costs (in the region of \$3000) were not insignificant, equating to a cost of \$250 per resident. It may be possible to substitute orthopedic fellows or staff in place of SPs in future iterations. However, given the importance of establishing competence in these areas by both the Royal College and the ACGME, this could be seen as a reasonable cost for training programs to bear on an annual basis.

The Royal College has a published handbook detailing assessment methods for the CanMEDS competencies.²⁷ It states that oral examinations and OSCEs are not well suited to evaluate the roles of manager and scholar. Other documents attest to the perceived difficulty with assessing the intrinsic roles, especially the role of health advocate.^{5,28,29} However, the reliability of the manager role in our study was high enough to be used in a high-stakes examination. While we cannot attest to the reliability of the health advocate and scholar roles owing to insufficient items, ANOVA testing demonstrated a significant ability to distinguish between residents of different PGY levels for both of these roles. We believe that the OSCE is a very appropriate means of assessment, as clinical scenarios that mimic real life encounters can be used.

An advantage of this type of OSCE is that both teaching (formative evaluation) and assessment (summative evaluation) can be incorporated. As noted by Zuckerman and colleagues,⁴ assessment motivates residents to learn important skills and is therefore a form of learning in itself. We believe that by exposing very junior residents (PGY0) to

scenarios they will likely soon encounter (complaints of delayed surgery, difficult interaction with operating room staff), learning opportunities can be created in an environment suitable for feedback and coaching.³⁰ Furthermore, by retesting mid-rank residents (PGY3), an assessment of their skills in each of the CanMEDS roles can be re-evaluated, and appropriate feedback can be provided. At our institution, a bank of multiple CanMEDS scenarios has been created; we believe that all residents will benefit from exposure to a CanMEDS OSCE twice in their training, once as a junior and once as a senior resident.

We are not aware of any OSCE designed to test only the intrinsic CanMEDS roles. While it is difficult to remove the medical expert role from such an examination, every effort was made to minimize scenarios dependent on orthopedic knowledge. For example, in the station focusing on the role of manager, residents were asked to manage an overbooked trauma list; some degree of orthopedic knowledge was required to know the urgency of each case, but residents were graded on their reasoning and on their ability to handle a phone call from a disgruntled relative. In the scholar station, residents were expected to know levels of evidence and how to perform database searches; in the needlestick case (professional role) residents were expected to know the immediate and delayed management of such an occurrence as well as the ethical principles involved regarding patient consent and notification of the appropriate monitoring bodies. For this reason, we do not believe that there were any major qualitative differences regarding the degree of core knowledge assessed in each station.

We were interested in obtaining concurrent validity; for example how could the station creator be certain that communication stations were truly assessing the communicator role. All case scenarios were based on real-life clinical situations, and adherent to role descriptions provided by the Royal College of Physicians and Surgeons Canada.^{1,6} Interestingly, no correlation was seen between the ITER role scores and the OSCE role scores, but a good correlation was seen with program director ratings of the roles. This suggests that ITERs are not a particularly effective form of assessment for the intrinsic roles.

Limitations

Limitations included our inability to comment on the reliability of the roles of scholar and health advocate, as only a single global rating was used for each of these roles — this will be remedied in the future. However, each of these roles was useful in distinguishing between different years of training. Objectivity may have been increased by the use of SPs or SHPs to provide global ratings of the residents, a method that has been used to good effect in the medical education literature, with evidence of good correlation between ratings completed by SPs and faculty

physicians.^{15,19,31} Importantly, the examiners will have known some residents and their PGY of training, raising the potential for bias. Examiners were asked to disregard the PGY level of the resident; however, it may be that the use of SPs' ratings will help to offset this risk. In this OSCE, neither station nor role weighting was used, as it was felt that each of the CanMEDS roles was equally important. Finally, it is not possible to know how this CanMEDS OSCE compares to a more traditional OSCE with incorporated assessment of CanMEDS roles within those stations; however, we have demonstrated a high degree of reliability or internal consistency, a measure that indicates an exam is performing well. It may be that the high degree of reliability seen in this CanMEDS OSCE may be a result of its narrow focus.

CONCLUSION

An OSCE designed to assess the intrinsic CanMEDS roles proved to be sufficiently valid and reliable for regular use in an orthopedic residency program.

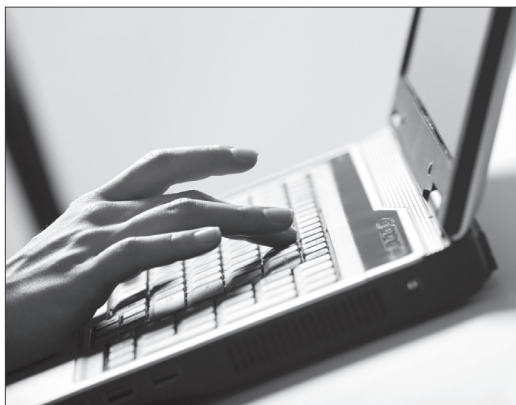
Competing interests: None declared.

Contributors: T. Dwyer, P. Ferguson, M.L. Murnaghan, B. Hodges and D. Ogilvie-Harris designed the study. T. Dwyer, D. Wasserstein, M. Nousiainen, P. Ferguson, V. Wadey, T. Leroux and D. Ogilvie-Harris acquired the data, which T. Dwyer, S. Glover-Takahashi, M. Hynes, J. Herold, J.L. Semple, B. Hodges and D. Ogilvie-Harris analyzed. T. Dwyer, S. Glover-Takahashi, M. Hynes, J. Herold, B. Hodges and D. Ogilvie-Harris wrote the article, which all authors reviewed and approved for publication.

References

1. Frank JR. The CanMEDS 2005 physician competency framework. Ottawa (ON): The Royal College of Physicians and Surgeons of Canada; 2005.
2. ACGME. Common program requirements: general competencies [website of ACGME]. Available: www.acgme.org/acgmeweb/Portals/0/PFAssets/ProgramRequirements/CPRs2013.pdf (accessed 2013 Mar. 25).
3. Jefferies A, Simmons B, Ng E, et al. Assessment of multiple physician competencies in postgraduate training: utility of the structured oral examination. *Adv Health Sci Educ Theory Pract* 2011;16:569-77.
4. Zuckerman JD, Holder JP, Mercuri JJ, et al. Teaching professionalism in orthopaedic surgery residency programs. *J Bone Joint Surg Am* 2012;94:e51.
5. Chou S, Cole G, McLaughlin K, et al. CanMEDS evaluation in Canadian postgraduate training programmes: tools used and programme director satisfaction. *Med Educ* 2008;42:879-86.
6. Jefferies A, Simmons B, Tabak D, et al. Using an objective structured clinical examination (OSCE) to assess multiple physician competencies in postgraduate training. *Med Teach* 2007;29:183-91.
7. Catton PTS, Rothman A. A guide to resident assessment for program directors. *Ann Roy Coll Phys Surg Can* 1997;403-9.
8. Schwartz RW, Witzke DB, Donnelly MB, et al. Assessing residents' clinical performance: cumulative results of a four-year study with the Objective Structured Clinical Examination. *Surgery* 1998;124:307-12.
9. Hodges B, Regehr G, Hanson M, et al. Validation of an objective structured clinical examination in psychiatry. *Acad Med* 1998;73:910-2.

10. Cohen R, Reznick RK, Taylor BR, et al. Reliability and validity of the objective structured clinical examination in assessing surgical residents. *Am J Surg* 1990;160:302-5.
11. O'Sullivan P, Chao S, Russell M, et al. Development and implementation of an objective structured clinical examination to provide formative feedback on communication and interpersonal skills in geriatric training. *J Am Geriatr Soc* 2008;56:1730-5.
12. Srinivasan J. Observing communication skills for informed consent: an examiner's experience. *Ann R Coll Physicians Surg Can* 1999;32:437-40.
13. Keely E, Myers K, Dojeiji S. Can written communication skills be tested in an objective structured clinical examination format? *Acad Med* 2002;77:82-6.
14. Yudkowsky R, Alseidi A, Cintron J. Beyond fulfilling the core competencies: an objective structured clinical examination to assess communication and interpersonal skills in a surgical residency. *Curr Surg* 2004;61:499-503.
15. Donnelly MB, Sloan D, Plymale M, et al. Assessment of residents' interpersonal skills by faculty proctors and standardized patients: a psychometric analysis. *Acad Med* 2000;75:S93-5.
16. Hodges B, McIlroy JH. Analytic global OSCE ratings are sensitive to level of training. *Med Educ* 2003;37:1012-6.
17. Hodges B, Turnbull J, Cohen R, et al. Evaluating communication skills in the OSCE format: reliability and generalizability. *Med Educ* 1996;30:38-43.
18. Warf BC, Donnelly MB, Schwartz RW, et al. The relative contributions of interpersonal and specific clinical skills to the perception of global clinical competence. *J Surg Res* 1999;86:17-23.
19. Colliver JA, Swartz MH, Robbs RS, et al. Relationship between clinical competence and interpersonal and communication skills in standardized-patient assessment. *Acad Med* 1999;74:271-4.
20. Sloan DA, Donnelly MB, Johnson SB, et al. Assessing surgical residents' and medical students' interpersonal skills. *J Surg Res* 1994;57:613-8.
21. Fliegel JE, Frohna JG, Mangrulkar RS. A computer-based OSCE station to measure competence in evidence-based medicine skills in medical students. *Acad Med* 2002;77:1157-8.
22. Schol S. A multiple-station test of the teaching skills of general practice preceptors in Flanders, Belgium. *Acad Med* 2001;76:176-80.
23. Altshuler L, Kachur E. A culture OSCE: teaching residents to bridge different worlds. *Acad Med* 2001;76:514.
24. Singer PA, Pellegrino ED, Siegler M. Clinical ethics revisited. *BMC Med Ethics* 2001;2:E1.
25. Probyn LF, Finley K. The CanMEDS objective structured clinical examination (OSCE): Evaluating outside of the box — an example from diagnostic radiology. Ottawa (ON): Royal College of Physicians and Surgeons of Canada; 2010. Available: www.royalcollege.ca/portal/page/portal/rc/common/documents/canmeds/whatworks/canmeds_probyn_finlay_osce_e.pdf.
26. Hilliard RIT, Tabak S.E. D. Use of an objective structured clinical examination as a certifying examination in pediatrics. *Ann R Coll Physicians Surg Can* 2000;33:222-8.
27. Bandiera GS, Sherbino J, Frank JR. The CanMEDS assessment tools handbook: an introductory guide to assessment methods for the CanMEDS competencies. Ottawa (ON): Royal College of Physicians and Surgeons of Canada; 2006.
28. Verma S, Flynn L, Seguin R. Faculty's and residents' perceptions of teaching and evaluating the role of health advocate: a study at one Canadian university. *Acad Med* 2005;80:103-8.
29. Murphy DJ, Bruce D, Eva KW. Workplace-based assessment for general practitioners: using stakeholder perception to aid blueprinting of an assessment battery. *Med Educ* 2008;42:96-103.
30. Duffy FD, Gordon GH, Whelan G, et al. Assessing competence in communication and interpersonal skills: the Kalamazoo II report. *Acad Med* 2004;79:495-507.
31. Cooper C, Mira M. Who should assess medical students' communication skills: their academic teachers or their patients? *Med Educ* 1998;32:419-21.



Canadian Journal of Surgery

**Online manuscript
submission and peer review**

AVAILABLE at

<http://mc.manuscriptcentral.com/cjs>