

Practice simulated office orals as a predictor of Certification examination performance in family medicine

Kendall Noel MDCM MEd CCFP FCFP Douglas Archibald PhD Carlos Brailovsky MD MA(Ed) MCFP(Hon)

Abstract

Objective To determine if performance on practice simulated office orals (SOOs) conducted during residency training could predict residents' performance on the SOO component of the College of Family Physicians of Canada's (CFPC's) final Certification examination.

Design Prospective cohort study.

Setting University of Ottawa in Ontario.

Participants Family medicine residents enrolled in the University of Ottawa's Family Medicine Residency program between July 1, 2012, and June 30, 2014, who were eligible to write the CFPC Certification examination in the spring of 2014 and who had participated in all 4 practice SOO examination sessions; 23 residents met these criteria.

Main outcome measures Scores on practice SOO sessions during fall 2012, spring 2013, fall 2013, and spring 2014; and the SOO component score on the spring 2014 administration of the CFPC Certification examination.

Results Weighted least squares regression analysis using the 4 practice SOO session scores significantly predicted the final Certification examination SOO score ($P < .05$), with an adjusted R^2 value of 0.29. Additional analysis revealed

that the mean scores for the cohort generated at each time point were statistically different from each other ($P < .001$) and that the relationship over time could be represented by either a linear relationship or a quadratic relationship. A generalizability study generated a relative generalizability coefficient of 0.63.

Conclusion Our results confirm the usefulness of practice SOOs as a progress test and demonstrate the feasibility of using them to predict final scores on the SOO component of the CFPC's Certification examination.

EDITOR'S KEY POINTS

- Simulated office orals (SOOs) are used by the College of Family Physicians of Canada to evaluate family medicine residents' readiness for clinical practice. It would be useful for residency programs to be able to use the practice examinations conducted during training to predict residents' performance on the College examination, in order to help identify those at risk of failure.
- Weighted least squares analysis, a generalizability study, and additional analyses all reported results that confirmed the usefulness of the SOO as a progress test. This study shows that it would be worthwhile for programs to formalize their practice SOO sessions and to conduct the analysis necessary to generate sound conclusions. Such analysis would allow programs to identify residents who might require greater assistance, which might range from more direct supervision to more robust remedial rotations.

This article has been peer reviewed.
Can Fam Physician 2017;63:299-305

Les pratiques d'entrevues médicales simulées comme prédicteur des résultats à l'examen de certification en médecine familiale

Kendall Noel MD CM MEd CCFP FCFP Douglas Archibald PhD Carlos Brailovsky MD MA(Ed) MCFP(Hon)

Résumé

Objectif Déterminer si les résultats aux entrevues médicales simulées (EMS) auxquelles les résidents sont soumis durant leur formation pourraient prédire la performance aux examens de même type utilisés par le Collège des médecins de famille du Canada (CMFC) à l'examen final de certification.

Type d'étude Étude de cohorte prospective.

Contexte L'Université d'Ottawa, en Ontario.

Participants Des résidents inscrits au programme de médecine familiale de l'Université d'Ottawa entre le 1^{er} juillet 2012 et le 30 juin 2014 qui étaient admissibles à l'examen de certification du CMFC au printemps 2014 et qui avaient participé aux 4 séances d'EMS; 23 résidents répondaient à ces critères.

Principaux paramètres à l'étude Les scores obtenus aux séances d'exercice des EMS de l'automne 2012, du printemps 2013, de l'automne 2013 et du printemps 2014; et les scores obtenus aux composantes EMS de l'examen de certification du CFPC du printemps 2014.

Résultats L'analyse de régression des moindres carrés pondérés effectuée avec les 4 séances de pratique d'EMS était un prédicteur significatif du score à l'examen final de certification ($P < .05$), avec une valeur ajustée de 0,29 pour le R². Les analyses additionnelles ont révélé que les scores moyens pour les cohortes constituées à chaque moment donné dans le temps étaient significativement différents les uns des autres ($P < .001$) et que leur évolution dans le temps pouvait correspondre à une relation linéaire ou à une relation quadratique. Une étude sur la possibilité de généraliser ces conclusions a donné un coefficient de généralisation de 0,63.

Conclusion Nos résultats confirment l'utilité des exercices d'EMS comme moyen de vérifier les progrès des résidents et qu'on peut utiliser leurs résultats pour prédire les résultats aux composantes EMS de l'examen final de certification du CMFC.

POINTS DE REPÈRE DU RÉDACTEUR

- Les entrevues médicales simulées (EMS) sont utilisées par le Collège des médecins de famille du Canada pour vérifier l'état de préparation à la pratique des résidents en médecine familiale. Les programmes de résidence auraient avantage à se servir de ce type d'examens durant la formation pour prédire la performance des résidents aux examens du Collège, ce qui permettrait d'identifier ceux qui sont à risque d'échec.
- L'analyse des moindres carrés pondérés, un type d'étude dont les résultats peuvent être généralisés, ainsi que d'autres analyses ont toutes donné des résultats qui confirment l'utilité des EMS pour vérifier les progrès accomplis. La présente étude montre que les programmes auraient avantage à officialiser leurs séances pratiques d'EMS et à effectuer les analyses nécessaires pour générer des conclusions valables. Les programmes pourraient ainsi identifier les résidents qui ont besoin d'une aide accrue allant d'une supervision plus directe à des stages de rattrapage plus soutenus.

Cet article a fait l'objet d'une révision par des pairs.
Can Fam Physician 2017;63:299-305

Shortly after the world was introduced to the use of standardized patients in medical examinations, through the seminal paper produced by Barrows and Abrahamson, the educational leadership at the College of Family Physicians of Canada (CFPC) began development of a performance examination that made use of this novel form of assessment.^{1,2} As the CFPC developed the examination, it hired Barrows himself and J.L. Maatsch, a contemporary who believed strongly in the use of simulated scenarios, for their help (P. Rainsberry, PhD, oral communication, 2015).³ The final product, a collection of novel multidimensional assessment tools, included 3 oral components in its 1969 inaugural session. One of these components, the simulated office oral (now colloquially known as the SOO), marked the first time in history that a standardized or programmed patient was used in a national certification examination.^{1,4,5}

Simulated office orals are structured oral examinations conducted with standardized patients; while they initially included laypeople playing the role of the patient, in 1984 the CFPC replaced laypeople with trained family medicine physicians, giving these new standardized patients the dual role of patient and SOO examiner.¹ This important aspect helped distinguish the SOO from the more popular structured oral standardized patient examination that was used then in North America—the objective structured clinical examination—by capitalizing on the theoretical symbiotic advantages seen when a physician examiner is allowed to make competency judgments and, at the same time, to personally experience the nonverbal communication (eg, eye contact) that plays out during a clinical encounter.⁶

The use of practice examinations to help prepare students across disciplines and levels of education has been studied extensively.⁷ Dotson's thesis paper provides an extensive review of the literature on the use of practice examinations, including an article describing the assessment-accuracy hypothesis.^{7,8} When applied to residency training, the assessment-accuracy hypothesis explains that better in-training assessment accuracy will allow residents to perform better on examinations, as they will be given the opportunity to modify their studying accordingly, well before challenging their high-stakes certification examination. While above-average students have been shown to be better at predicting their test performance than below-average students are, the value of an external assessment of a resident's abilities before writing a high-stakes examination cannot be overemphasized, even if it can be assumed that most medical residents were at one time above-average undergraduate students.⁹

For American programs accredited under the auspices of the Accreditation Council on Graduate Medical Education, the predominant method for in-training

assessment of medical knowledge has been the specialty-specific in-training examination (ITE), a written examination. While there is some debate surrounding the ability of ITEs to predict those who will ultimately fail their board examinations, most studies acknowledge the advantage of such examinations, namely providing residency directors with the ability to clearly identify those at risk of failure.^{10,11} As there is no formal Canadian equivalent to the ITE in family medicine, practice SOO sessions for residents represented an interesting opportunity for study.

In Canada, the use of practice SOOs as a formative evaluation tool is common for most residency programs.¹² The CFPC has previously established the relevance of the SOOs themselves and reported that they have high content, construct, and face validity, with reasonable reliability.⁵ Furthermore, the CFPC believes that the SOOs are the best method for the College to assess a candidate's ability to establish an effective patient-doctor relationship.⁵ While the qualitative feedback from practice SOO examiners should in theory assist residents in preparing for their final Certification examination, few programs go on to convert the results into a numeric value in a manner consistent with the College's analysis of the Certification examination results.¹²

Recognizing that it might be easy to identify a resident who has performed extremely poorly on a practice SOO, the real power of a numerical analysis is in the identification of borderline candidates. This article presents the results of a study that explored the ability of practice SOOs to predict the scores generated on the SOO component of the final Certification examination.

METHODS

Study participants

During the fall of 2013 and the spring of 2014, family medicine residents at the University of Ottawa in Ontario were invited to participate in this study. The protocol was reviewed and ethics approval was provided by both the Bruyère Research Board and the Ottawa Hospital Research Institute Research Ethics Board. Residents enrolled in the program between July 1, 2012, and June 30, 2014, who were eligible to write the CFPC Certification examination in the spring of 2014 are included in the data presented.

A total of 65 residents were eligible to be included in the study. Forty-four (67.7%) of these residents signed up for the study, of which 23 (52.3%) attended all 4 practice SOOs. All 23 residents were Canadian graduates, 14 (60.9%) were female, and all were successful on the SOO component of their Certification examination in the spring of 2014.

Testing instruments and session format

Two practice SOO sessions are organized each year by the Department of Family Medicine at the University of Ottawa for its residents. The sessions are administered during the fall (October) and spring (March) of each academic year, resulting in a total of 4 administrations during the 2-year family medicine curriculum. During each session the residents are provided with 2 cases representing common types of problems that they might encounter during a typical office day. The cases are selected from a bank of released SOOs made available to the training programs by the CFPC. The first author (K.N.) chose the selected cases after reviewing them for content and psychometric properties. The SOOs are generated by the CFPC Committee on Examinations, following their established blueprinting process, and are aimed at the level of a family medicine graduate ready for independent practice. **Table 1** provides a list of the medical problems presented in the SOO cases used for the study.

Faculty members at the University of Ottawa volunteered to act as examiners during the practice sessions. Examiners were required to attend a training session held the evening before the practice SOO sessions in order to ensure standardization of their role-playing and marking responsibilities.

Each SOO had its own specific rubric, detailing how it was to be scored. For each SOO, marks were awarded in 6 areas: identification of the first problem; identification of the second problem; identification of the social and developmental context unique to the patient in question; management of the first problem; management of the second problem; and an overall mark for how the candidate conducted the interview. The rubric is divided into a left-hand-side score and a right-hand-side score for the first 5 areas. The left-hand-side score is generated using a graded checklist looking at aspects more commonly associated with the traditional voice of medicine. The right-hand side of the rubric represents the voice of the real world and a measure of the candidate's ability to be patient-centred; it is more subjective, with a listing of key features to be identified or demonstrated by the candidate.¹³ A more detailed description of the SOO rubric has been previously published.⁵

Study setting

The study took place within the context of the University of Ottawa's family medicine training program. The program includes 7 teaching sites, including rural-based community practices, urban-based community practices, and urban-based teaching units. All residents at the University of Ottawa were provided an introductory lecture, which reviewed the patient-centred clinical method as it applied to the SOOs, within the first 3 months of their first year. Some units provide their residents with additional sessions in their second year. In addition to the 8 official practice sessions, some residents had additional opportunities to practise SOOs with colleagues or faculty members in preparation for their examinations. These were generally conducted just before and after the last practice session in the spring of their second year.

Marks for the SOO component

The marks for the SOO component of the CFPC's Certification examination were obtained from the College for those residents who had enrolled in the study. As the College was in a period of transition, the SOO data had to be extracted from the clinical skills examination mark, which included marks for both the SOOs and the Medical Council of Canada's objective structured clinical examination. A member of our research team (C.B.) was responsible for sorting through the College data and provided us with the required Certification examination SOO scores. The data were then merged with the data obtained during the practice SOO sessions.

Analysis

Each SOO was scored according to its rubric to generate a mark. For each session (fall 2012, spring 2013, fall 2013, and spring 2014), the scores for the 2 SOOs were tallied and reported as a percentage score.

A generalizability (G) study was conducted using the software EduG, version 6.0-e, with the purpose of determining whether the SOO instruments were able to reliably differentiate among the residents.¹⁴ The study included 3 facets: residents, time, and the SOO items. The resident facet was crossed with the other 2 facets, as each resident challenged all 8 SOOs and each resident attended all 4 practice SOO sessions. The SOO item facet

Table 1. Problems or medical conditions presented in the SOOs for this study

SOO PRACTICE SESSIONS	CASE DESCRIPTION FOR SOO A	CASE DESCRIPTION FOR SOO B
Fall 2012	Patient with PTSD who is also concerned about his son's headache	Patient with nausea of pregnancy who is responsible for her dying father
Spring 2013	Patient with cluster headaches and paranoid schizophrenia	Patient with anogenital warts and sexual assault experience
Fall 2013	Patient with chronic disease and PTSD	Patient with smoking cessation and memory loss
Spring 2014	Post-MI patient with depression	Patient with COPD experiencing elder abuse

COPD—chronic obstructive pulmonary disease, MI—myocardial infarction, PTSD—posttraumatic stress disorder, SOO—simulated office oral.

was nested in the time facet, as each SOO item took place at a certain time point. Given that it was possible to imagine an infinite number of residents and an infinite number of cases, the estimation design for the study set both of these facets to infinite random. The time facet was set at a fixed number of 4. As per tradition, the primary observation design was set with residents as the object of differentiation and the SOO items nested in time as the object of measurement (expressed as $R \times S:T$ [residents crossed by SOO items nested in time]).¹⁵

A weighted least squares regression analysis was conducted to determine if the Certification SOO score could be predicted by fall 2012, spring 2013, fall 2013, and spring 2014 SOO marks. Finally, the results of a repeated-measures analysis of the data for those residents who participated in all 4 time points and the final Certification examination were reported to further demonstrate the feasibility of a SOO being used as a progress test.

The raw data were entered in Excel and then imported into SPSS, version 21, for all statistical calculations, excluding the G study.

RESULTS

Generalizability study

The first G study using residents as the focus of the differentiation demonstrated that 11.6% of the variance was owing to the differentiation facet (residents), 27.5% of the variance was owing to the timing of the practice sessions, and 1.7% of the variance was owing to the SOO items nested in time or, alternatively, owing to the difficulty of the SOO cases themselves. Most of the variance (53.5%) was owing to the interaction between residents and the SOO items nested in time plus any residual error. The resultant relative G coefficient for the measurement design of $R \times S:T$ using the data obtained for the 8 SOOs conducted over the 2 years of the study was 0.63. This value falls within the specifications noted as being the minimum necessary for a test that measures a multidimensional construct such as clinical reasoning.¹⁶ The facets and ANOVA (analysis of variance) results for the SOO G study are summarized in **Tables 2** and **3**, respectively.

Weighted least squares regression

Multiple regression was conducted to determine the best linear combination of practice SOO marks (fall 2012, spring 2013, fall 2013, and spring 2014). The means, standard deviations, and intercorrelations can be found in **Table 4**.

This combination of variables significantly predicted the spring Certification examination SOO score ($F_{4,22}=3.26$, $P<.05$) with one of the variables, fall 2012, contributing significantly to predicting the Certification SOO mark. The adjusted R^2 was 0.29, indicating that 29% of

Table 2. Problems or medical conditions presented in the SOOs for this study

FACETS	LABELS	LEVELS	UNIVERSE
Residents	R	23	INF
Time	T	4	4
SOOs nested in time	S:T	2	INF

INF—infinite, SOO—simulated office oral.

the variance in the SOO score was explained by the model. The results are summarized in **Table 5**.

Repeated-measures analysis

A repeated-measures ANOVA was conducted on the scores of the 23 residents who attended all 4 practice SOO examination sessions and the Certification examination's SOO component. Mauchly test for sphericity was not significant ($P>.05$), indicating that the data met the criteria for using a univariate approach to repeated-measures ANOVA. The ANOVA analysis indicated that there were differences between the resident scores ($F_{4,88}=23.85$, $P<.001$, $\eta^2=0.52$). Polynomial contrasts analysis indicated that there was progression and that this could be best represented either by a linear relationship ($F_{1,22}=137.09$, $P<.001$, $\eta^2=0.86$) or by a quadratic relationship ($F_{1,22}=19.13$, $P<.001$, $\eta^2=0.47$).

DISCUSSION

The purpose of the study was to determine if the scores generated on practice SOOs conducted during residency could predict how residents would perform on the SOO component of the Certification examination, thereby justifying the time and energy involved in running practice SOO sessions.

The results of our multiple regression analysis using all 4 practice SOO marks was statistically significant, accounting for 29% of the variance as calculated by the adjusted R^2 reported for the regression analysis. This model compares favourably with 2 published studies looking at predictors related to the SOO marks.^{17,18}

In the first study, the authors generated 2 models: one based on professional experience factors (internship, previous residency, professional experience, and research experience) and the other based on demographic characteristics (country of birth, human development index value, age, years since graduation, and first language).¹⁷ They found that professional experience and demographic characteristics explained 7% and 15% of the variance, respectively, for the marks generated by international medical graduates who were residency trained.¹⁷

In the second study, the authors sought to look at factors that would predict how practice-eligible

Table 3. The ANOVA (analysis of variance) results for the SOO generalizability study

SOURCE	SS	df	MS	COMPONENTS			
				RANDOM	MIXED	CORRECTED	PERCENTAGE*
R	1124.15	22	51.10	3.54	4.05	4.05	11.6
T	1879.35	3	626.45	12.83	12.83	9.62	27.5
S:T	129.52	4	32.38	0.59	0.59	0.59	1.7
R×T	1502.15	66	22.76	2.02	2.02	2.02	5.8
R×S:T	1646.48	88	18.71	18.71	18.71	18.71	53.5

MS—mean square, R—residents, R×S:T—residents crossed by SOOs nested in time, R×T—residents crossed by time, SOO—simulated office oral, SS—sum of squares, S:T—SOOs nested in time, T—time.

*Percentages do not add to 100 owing to rounding.

Table 4. Descriptive statistics and the Pearson correlations of variables used in regression model

VARIABLES	MEAN (SD) MARK, %	CORRELATIONS				
		VARIABLE 1	VARIABLE 2	VARIABLE 3	VARIABLE 4	VARIABLE 5
Certification examination SOO mark (variable 1)	66.13 (1.47)	1.0	0.57*	0.38 [†]	-0.11	0.21
Practice SOO marks from the following sessions:						
• Fall 2012 (variable 2)	49.86 (2.21)	0.57*	1.0	0.26	-0.10	0.15
• Spring 2013 (variable 3)	52.85 (2.66)	0.38 [†]	0.26	1.0	0.36 [†]	0.61*
• Fall 2013 (variable 4)	70.98 (2.28)	-0.11	-0.10	0.36 [†]	1.0	-0.06
• Spring 2014 (variable 5)	69.81 (2.15)	0.21	0.15	0.61*	-0.06	1.0

SOO—simulated office oral.

* $P < .01$.

[†] $P < .05$.

Table 5. Problems or medical conditions presented in the SOOs for this study

VARIABLES	β COEFFICIENT	STANDARD ERROR (β)	β
Fall 2012 SOO score	0.30	0.13	0.46*
Spring 2013 SOO score	0.24	0.15	0.43
Fall 2013 SOO score	-0.14	0.14	-0.22
Spring 2014 SOO score	-0.09	0.17	-0.13

SOO—simulated office oral.

* $P < .05$.

physicians would score on the examination and they found that using a model that included the most recent Medical Council of Canada Equivalency Examination and the Medical Council of Canada Qualifying Examination part 1 scores, sex, age in years at the start of the practice-ready assessment process, years since obtaining the medical degree, and language in which the medical degree was completed, that they were able to account for roughly 6% of the variance.¹⁸

Our G study provides further assurances. The G coefficient (relative) for the R×S:T measurement design equals 0.63 and, while low, it is compatible with reliability studies done on SOOs in the past and is acceptable given the multidimensional construct that the examination seeks to measure.⁵

Finally, furthering our work on the use of SOOs as a progress test, we report the results of the repeated-measures

analysis conducted on the data, including the results of the Certification examination SOO.¹⁹ The reported results provided further confirmation of the usefulness of the SOO as a progress test.

Limitations

Our study includes a few limitations that are worth mentioning. First, it is a relatively small study focused on the observations of one residency program. That said, the variability of experiences that residents trained in our program encounter is large and reflective of the experiences that can be had by residents in the wider family medicine community nationally. Second, our numbers are small; however, statistical significance was demonstrated, and more important the effect size, as represented by our adjusted R^2 value, was comparable to that seen in the literature.^{17,18}


Reflections and unanswered questions

Entering the third year of our study and preparing to collect second-year resident data from our second cohort has allowed time to reflect on a number of questions. Could breaking down the SOO marks provide greater insight into the various elements required for clinical reasoning and, if so, would that mean that practice SOOs could provide an opportunity to “diagnose” problems in clinical reasoning?²⁰ In a continuing effort to understand why some strong residents and practice-eligible candidates fail, one must ask, is there something intrinsic to the SOO instrument, specifically its rubric, that inadvertently disadvantages individuals with more experience or clinical acumen?²¹⁻²³ Finally, what accounts for the reported higher odds of passing the SOO component on the second try seen for practice-ready candidates, and how might the CFPC use this knowledge to guide future practice-ready candidates?¹⁸

As a result of these reflections and unanswered questions, we encourage other family medicine residency programs to collaboratively begin the task of looking at these and other issues and to address them in the literature.

Conclusion

Practice SOO examinations during family medicine residency are common and yet their formal use as a predictor of Certification examination performance has not been described in the literature. Our study provides evidence that it would be worthwhile for programs to formalize their practice SOO sessions and to conduct the analysis necessary to generate sound conclusions, including the generation of risk assessment plots.¹⁹ Such analysis would allow programs to identify residents who might require greater assistance early on, which might range from something as simple as more direct supervision, to a more robust, full-fledged remedial rotation.

Finally, now that we have established that the “laboratory testing” of the SOO in a residency program provides a reasonable approximation of the Certification examination process, we hope that the research community will be poised to begin the work necessary to better understand the role of the SOO in the assessment of clinical reasoning skills. 

Dr Noel is Assistant Professor in the Department of Family Medicine at the University of Ottawa in Ontario, a doctoral student in the PhD in Family Medicine program at Western University in London, Ont, and a member of the College of Family Physicians of Canada's Board of Examiners Committee. **Dr Archibald** is Assistant Professor and Education Researcher in the Department of Family Medicine at the University of Ottawa. **Dr Brailovsky** is Emeritus Professor at Laval University in Quebec city, Que, and is a consultant in testing and assessment at the College of Family Physicians of Canada in Mississauga, Ont.

Acknowledgment

This project was funded by Program and Innovation in Medical Education grants courtesy of the University of Ottawa's Department of Family Medicine, as overseen by the C.T. Lamont Primary Health Care Research Centre at the Bruyère Research Institute.

Contributors

All authors contributed to the concept and design of the study; data gathering, analysis, and interpretation; formulation of the questions; and preparing the manuscript for submission.

Competing interests

Drs Noel and Brailovsky are both members of the College of Family Physicians of Canada's Board of Examiners Committee. **Dr Brailovsky** sits on the committee as a psychometric consultant, while **Dr Noel** is a voting member of the board.

Correspondence

Dr Kendall Noel; e-mail knoel100@uottawa.ca

References

- Grand'Maison P, Lescop J, Brailovsky CA. Canadian experience with structured clinical examinations. *CMAJ* 1993;148(9):1573-6.
- Barrows HS, Abrahamson S. The programmed patient: a technique for appraising student performance in clinical neurology. *J Med Educ* 1964;39:802-5.
- Maatsch JL. Assessment of clinical competence on the Emergency Medicine Specialty Certification Examination: the validity of examiner ratings of simulated clinical encounters. *Ann Emerg Med* 1981;10(10):504-7.
- Lamont CT, Hennen BK. The use of simulated patients in a certification examination in family medicine. *J Med Educ* 1972;47(10):789-95.
- Brown JB, Handfield-Jones R, Rainsberry P, Brailovsky CA. Certification Examination of the College of Family Physicians of Canada. Part 4: simulated office orals. *Can Fam Physician* 1996;42:1539-42, 1545, 1547-8.
- Swanson DB, van der Vleuten CP. Assessment of clinical skills with standardized patients: state of the art revisited. *Teach Learn Med* 2013;25(Suppl 1):S17-25.
- Dotson WH. *Investigating the variables in a mock exam study session designed to improve student exam performance in an undergraduate behavior modification and therapy course* [dissertation]. Lawrence, KS: University of Kansas; 2010.
- Balch WR. Practice versus review examinations and final exam performance. *Teach Psychol* 1998;25(3):181-5.
- Balch WR. Effect of class standing on students' predictions of their final exam scores. *Teach Psychol* 1992;19(3):136-41.
- Withiam-Leitch M, Olawaiye A. Resident performance on the in-training and board examinations in obstetrics and gynecology: implications for the ACGME Outcome Project. *Teach Learn Med* 2008;20(2):136-42.
- Corneille MG, Willis R, Stewart RM, Dent DL. Performance on brief practice examination identifies residents at risk for poor ABSITE and ABS qualifying examination performance. *J Surg Educ* 2011;68(3):246-9. Epub 2011 Feb 23.
- Greenberg G, Bradel T, Ganshorn K, Mahood S, Zagorski C, Lawrence K. Short report: preparing for simulated office orals. Survey of practices in 16 family medicine departments. *Can Fam Physician* 2002;48:745-7.
- Mishler EG. *The discourse of medicine. Dialectics of medical interviews*. Westport, CT: Greenwood Publishing Group; 1984.
- Cardinet J, Johnson S, Pini G. *Applying generalizability theory using EduG*. New York, NY: Routledge; 2010.
- Cardinet J, Tourneur Y, Allal L. The symmetry of generalizability theory: applications to educational measurement. *J Educ Meas* 1976;13(2):119-35.
- Brailovsky CA, Grand'Maison P, Lescop J. A large-scale multicenter objective structured clinical examination for licensure. *Acad Med* 1992;67(10 Suppl):S37-9.
- Schabert I, Mercuri M, Grierson LE. Predicting international medical graduate success on college certification examinations. Responding to the Thomson and Cohl judicial report on IMG selection. *Can Fam Physician* 2014;60:e478-84. Available from: www.cfp.ca/content/60/10/e478.full.pdf+html. Accessed 2017 Feb 24.
- De Champlain AF, Streefkerk C, Tian F, Roy M, Qin S, Brailovsky C, et al. *Predicting family medicine specialty certification status using standardized measures for a sample of international medical graduates engaged in a practice-ready assessment pathway to provisional licensure*. Ottawa, ON: Medical Council of Canada. Available from: <http://mcc.ca/wp-content/uploads/reports-predictive-validity-journal-article.pdf>. Accessed 2017 Feb 24.
- Noel K, Archibald D, Brailovsky C, Mautbur A. Progress testing in family medicine—a novel use for simulated office oral examinations. *Med Teach* 2016;38(4):364-8. Epub 2015 May 13.
- Audétat M, Lubarsky S, Blais J, Charlin B. Clinical reasoning: where do we stand on identifying and remediating difficulties? *Creat Educ* 2013;4(6):42-8.
- Terry R, Hiestler E, James GD. The use of standardized patients to evaluate family medicine resident decision making. *Fam Med* 2007;39(4):261-5.
- Schmidt HG, Boshuizen HP. On the origin of intermediate effects in clinical case recall. *Mem Cognit* 1993;21(3):338-51.
- Rikers RM, Schmidt HG, Boshuizen HP. Knowledge encapsulation and the intermediate effect. *Contemp Educ Psychol* 2000;25(2):150-66.