81: Short Communications: Assessment: Clinical

Location: Rio, 2nd Floor, CCB
Date: Tuesday 28th August
Time: 1400-1530 hrs

81 (3722)
Introducing KAANBAL for clinical competence assessment: A digital mobile tool

Authors
Cesar Alberto Lucio Ramirez, Tecnologico de Monterrey School of Medicine and Health Sciences, Monterrey, Mexico
Juan Pablo Nigenda Alvarez, Tecnologico de Monterrey School of Medicine and Health Sciences, Monterrey, Mexico
Silvia Lizett Olivares Olivares, Tecnologico de Monterrey School of Medicine and Health Sciences, Monterrey, Mexico
Mildred Vanessa Lopez Cabrera, Tecnologico de Monterrey School of Medicine and Health Sciences, Monterrey, Mexico
Jose Antonio Diaz Elizondo, Tecnologico de Monterrey School of Medicine and Health Sciences, Monterrey, Mexico

Presenter:
Cesar Alberto Lucio Ramirez, Tecnologico de Monterrey School of Medicine and Health Sciences, Monterrey, Mexico

Background: Medical schools and healthcare institutions require evidence from students' clinical competence learning and progress. A single evaluation based on one particular observation does not guarantee timely feedback for learners, faculty and authorities for continuous improvement. The objective of this project was to design and implement a mobile system to assess clinical competence progress for medical students during several clinical clerkship rotations.

Method: A mobile system was constructed with local software developers. Before starting its design, medical specialists from several disciplines were invited to participate on five focus groups to gather their impressions and needs regarding the clinical competence assessment. The most popular problems described by participants were: a) one single final grade lacking feedback for students' improvement, b) absence of objectiveness from examiners, and c) perfect scores to every student given only upon professional behavior. An app called KAANBAL was designed to be available for most devices. It includes customizable rubrics that allows real time feedback and full reports for both students and faculty. This app allows multiple assessments from different examiners across time. Students are invited to assess faculty performance. Before implementation, a faculty development training workshop was delivered for clinical educators to instruct them about medical assessment basics and the use of the mobile tool.

Results: A total of 2,958 clinical skills assessments were done in a 12-month period to assess 157 students. Students went from receiving on average 8 assessments to 18 per year. A survey for the 250 faculty that participated showed that 88% considered KAANBAL user friendly, 73% appraised feedback, 75% reported student engagement and 87% of faculty agreed that this system helped them improve their clinical teaching.

Discussion & Conclusion: Increasing the number of clinical skills assessments improves reliability of the scores and promotes learning for better patient care. KAANBAL has the potential to encourage continuous improvement on performance applying workplace based assessments. Simplification and automatization of the clinical skills assessment process can lead to more active participation. Medical schools and health care institutions have timely information for decision making.

812 (2252)
Does a clinical skills assessment practice session improve low achieving students' performance on the USA Medical Examination® (USMLE) Step 2 Clinical Skills (CS) Examination?

Authors
Kimberly Kirkland, American University of the Caribbean School of Medicine, Sint Maarten, Netherlands Antilles
Joy Checa, American University of the Caribbean School of Medicine, Sint Maarten, Netherlands Antilles
Robert Hecht, American University of the Caribbean School of Medicine, Sint Maarten, Netherlands Antilles
Steve Ash, American University of the Caribbean School of Medicine, Sint Maarten, Netherlands Antilles
William Behrns, American University of the Caribbean School of Medicine, Sint Maarten, Netherlands Antilles
Julie Taylor, American University of the Caribbean School of Medicine, Sint Maarten, Netherlands Antilles

Presenter:
Kimberly Kirkland, American University of the Caribbean School of Medicine, Sint Maarten, Netherlands Antilles

Background: Lower academic performance and lower USMLE Step 1 scores are associated with poor USMLE Step 2 CS outcomes. At American University of the Caribbean School of Medicine (AUC), passing an optional 8-station practice Clinical Skills Assessment (CSA) is predictive of passing the CS. In this study, we focus on lower achieving students to determine if their performance on the CS was positively impacted by completion of the CSA.

Method: At AUC, 1339 students between March 2014 and September 2017 were categorized as “high achieving” or “low achieving” based on their performance on the USMLE Step 1 (< or > 215) or on GPA (< or > 83.8). A logistic regression model was used. Categorical variable of interest was the CS outcome; independent variables were CSA attendance and performance, GPA, and Step 1 score.

Results: Performance on the CS of a cohort of 1339 students over a three-year period was analysed. Low academic performance and low Step 1 performance were found to be independently predictive of poor CS performance (P values: <0.001 and 0.002 respectively). Although the CSA was not found to be statistically significant for students with low achieving GPAs, it was approaching significance for students with low achieving Step 1 scores (P value 0.058). It’s likely that more data could prove this effect to be statistically significant.
Clinical deans need objective guidance to drive investment of resources to improve medical students’ success.

Conclusions: At our institution as at others, low academic and USMLE Step 1 performance is associated with poor outcomes on the CS. While a CSA practice session seems to have a positive impact on CS performance in low achieving Step 1 test-takers, the results were not significant. A student taking the CS soon after completing CSA may see a higher effect than a student with several months between the exams; follow-up analysis may be valuable to determine the impact of timing of the CSA on CS performance. Further investigation is needed before resources are invested in mandating the CSA for low achieving students.

8I3 (1923)

GOSCE - A method for training and evaluating clinical competence

Authors
Mahdi Ghassabi Chorsi
Mahmoud Mansouri
Hamid Reza Baradaran

Presenter:
Mahdi Ghassabi Chorsi, Iran University of Medical Sciences, Tehran, Iran

Background: The goals of clinical education in basic medical education are to obtain the minimum required skills for the graduates. The clinical competency exam is designed as a tool for evaluating these learning minima. There are various tools for evaluating these learning minima; among which GOSCE is very important. The purpose of this paper is to describe the features and benefits of GOSCE.

Method: This study was a kind of review. In order to collect information, articles from Science Direct, Pub Med, SID, Iran doc, ProQuest, Medline published in both Persian and English languages were used without time limit. For this purpose, the hierarchical search method and the key words of the clinical competency test, OSCE and GOSCE were used. From the selected articles, 26 articles were included.

Results & Conclusions: GOSCE is an effective and learner-centered approach covering multiple goals of communication skills training, clinical reasoning, self-assessment and feedback in a teaching-based environment. It has a positive effect on the attitude toward learning communication skills and self-learning in relation to the clinical environment. Also, learners are evaluated through feedback by counterparts, patients, and faculty members. GOSCE can be used as a suitable method for assessing clinical competence. This method has advantages such as increased learning motivation, self-assessment, feedback from multiple sources, automatic readiness, self-confidence, good patient relationship, low cost, and resources. It is suggested to replace OSCE by GOSCE.

Medical Sciences, Clinical Competency, OSCE and GOSCE

8I4 (829)

Complex Intrinsic Skill Competencies: A Fit-for-Purpose Multiple Component Assessment Tool

Authors
Debra Sibbald, Touchstone Institute, Toronto, Canada
Sandra Monteiro, Touchstone Institute, Toronto, Canada

Presenter:
Debra Sibbald, Touchstone Institute, Toronto, Canada

Background: Intrinsic skill (non-expert) competencies are important capacities for health care professionals, but they cannot be measured using traditional methods. Such interdependent abilities are often demonstrated through integrated patterns of behaviours. Interview formats are common but less reliable in predicting performance and are limited in uncovering the interplay of skills. Few studies have examined the utility of a multi-perspective approach combining validated assessment approaches to measure intrinsic skills.

Method: An innovative screening tool was designed at Touchstone Institute to evaluate non-expert CanMEDS roles in practice-ready family practitioners using multiple components of standardized performance and reflective assessments. The performance component consists of eight structured short stations. In seven, candidates respond to objective and reflective questions related to challenging scenarios assessing interpersonal, cognitive and decision-making skills. The eighth station uses an interview format. On all stations, a physician, nurse and standardized client rate each candidate using numeric global ratings for overall performance and entrustment. Raters also record written comments to capture qualitative assessments. The second component is a tablet questionnaire combining two validated instruments measuring personal reflection ability and insights: the Groningen Reflection Ability Scale (GRAS) and the Self-Reflection and Insight Scale (SRIS). A pilot test was conducted as a validation study, comparing Canadian trained family medicine practitioners to residents. Scores were evaluated for inter-rater reliability, and internal consistency.

Results: Internal consistency was high (α = 0.93) with acceptable means for overall and for entrustment scales (3.5/5). There were no meaningful differences between raters, as scores were highly correlated across all 3 raters (r > 0.7). Written comments were infrequent and primarily highlighted unprofessional behaviours. These candidates also received lower scores. The self-reflection ability score was not correlated with overall performance on the 8 stations (r = 0.039), suggesting unique constructs.

Conclusion: This design shows promise as a screening assessment, providing evidence of complex competencies derived through multiple measures and multiple perspectives. This innovative tool of pattern-based intrinsic skills offers rich, multi-disciplinary assessments of intrinsic skills and will be implemented in the selection of practice-ready international family physicians.
815 (1706)
Setting conjunctive standards in performance assessments: exploring the why and the how

Authors
Matt Homer
Dave Swanson
Godfrey Pell

Presenter:
Matt Homer, University of Leeds, UK

Background: Many institutions require students to achieve a minimum number of OSCE stations passed (MNSP) in addition to the aggregate pass mark. The rationale is that such conjunctive standards prevent excessive degrees of compensation across an assessment (i.e. students doing very well on a few stations, but poorly on many others). However, in the literature, we find there is generally a lack of underpinning empirical and theoretical evidence to support this approach, and the common practice of pre-determining a MNSP arguably lacks defensibility.

Method: We have surveyed the relevant literature to consider the psychometric arguments for and against such conjunctive standards. We have also developed a methodology for the setting of a post hoc conjunctive standard (MNSP) based on assessment outcomes. To do this we have used a methodology that is similar to that used in borderline regression standard setting - we use logistic regression to predict probabilities of passing stations based on global grades, and then aggregate these for the borderline group across the OSCE.

Results: There are psychometric arguments for and against the use of conjunctive standards in OSCES, but from a pragmatic point of view it is clear that many stakeholders approve and value this approach. Our work suggests that there is some variation in an empirically set number-of-stations-passed standard, but that few students fail to reach this standard in OSCEs.

Discussion: We have shown that it is possible to develop methods for measuring levels of compensation, and for producing a defensible MNSP standard. However, we find that the prevalence of excessive compensation is low, with few students failing based on the empirically set conjunctive standard. The underlying assumptions employed, and the implications for assessment policy of these findings, will be discussed.

Conclusion: Defensibility of decision-making in high stakes assessments is an absolute necessity. Ideally, conjunctive standards need to be justified alongside other assessment polices such the minimising of false positives via SEMs. However, taking into account wider policy issues, stakeholder views and the underlying methodological assumptions are also essential when doing this.

816 (2730)
A Study on Standard-Setting for the Application of Criterion-referenced Evaluation in Core Clinical Clerkship Summative Evaluation

Authors
Na Jin Kim, The Catholic University of Korea, College of Medicine, Seoul, South Korea
Young-Min Kim, The Catholic University of Korea, College of Medicine, Seoul, South Korea
Jaseong Koo, The Catholic University of Korea, College of Medicine, Seoul, South Korea
Sung A Lee, The Catholic University of Korea, College of Medicine, Seoul, South Korea
Youngjung Kim, The Catholic University of Korea, College of Medicine, Seoul, South Korea

Presenter:
Na Jin Kim, The Catholic University of Korea, College of Medicine, Seoul, South Korea

Background: Evaluation of clinical skill performance ability conducted in medical colleges of Korea is mostly based on norm-referenced scores to give a pass/ non-pass. This study recognizes the issues with clinical skill evaluation as a form of norm-referenced evaluation, and seeks to apply criterion-referenced evaluation in core clinical clerkship summative evaluation.

Method: With 96 medical students enrolled in their third year as of end of semester in 2017 as subjects, 6 CPX and 6 OSCE stations were conducted. 6 panelists were selected and trained individually. Applying the Modified Angoff and the Hofstee methods, a cut score was calculated and the process and results were both evaluated. It was reviewed internal, external and procedural validity of the score in terms of its feasibility.

Results: The results showed that CPX cut scores for cases ranged 43.8~65.3 by Angoff or 41.0~68.0 by Hofstee. The OSCE scores ranged 69.6~77.5 by Angoff and 66.0~89.0 by Hofstee. Intrapanelist consistency showed an overall high at .75(Kappa) or higher, and Interpanelist consistency was also high at .60 or higher. Inter-method consistency was high at .70. A questionnaire conducted to evaluate the procedural criteria found that the degree of understanding theoretical concepts and principles were relatively high, as well as validity of process. Moreover, the reasonableness of cut score was rated relatively positively. Overall, cut scores were lower or of similar level when the Angoff method was used, as opposed to the Hofstee method. To compile conflicting results into one criterion, the final cut score was set to be the mean value of the two methods, and the total score of 900(CPX 600/OSCE 300) was taken into account. This led to a non-pass rate of 29.1%. Further research is needed to: 1. A study on whether the same standard setting results should be applied in every year; 2. A study on predictive validity.

Conclusion: While comparative studies on standard setting using traditional methods and compromised methods are rare, based on the findings of this study, an appropriate method for establishing criterion for the evaluation of clinical skill performance ability was suggested.