ORIGINAL REPORTS

Development and Validation of a Tool to Evaluate the Evolution of Clinical Reasoning in Trauma Using Virtual Patients

David Fleiszer, MD,[*,†,1] Michael L. Hoover, PhD,[‡,2] Nancy Posel, PhD,[*,§] Tarek Razek, MD,[*,‖,3] and Simon Bergman, MD[*,¶,#,4]

[*]Faculty of Medicine, McGill University, Montreal, Quebec, Canada; [†]Departments of Surgery and Oncology, McGill University Health Center, Montreal, Quebec, Canada; [‡]Department of Educational and Counselling Psychology, McGill University, Montreal, Quebec, Canada; [§]McGill University Health Center, Montreal, Quebec, Canada; [‖]Montreal General Hospital, Montreal, Quebec, Canada; [¶]Department of Surgery, Jewish General Hospital, McGill University, Montreal, Quebec, Canada; and [#]Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec, Canada

CONTEXT: Undergraduate medical students at a large academic trauma center are required to manage a series of online virtual trauma patients as a mandatory exercise during their surgical rotation.

PURPOSE: Clinical reasoning during undergraduate medical education can be difficult to assess. The purpose of the study was to determine whether we could use components of the students' virtual patient management to measure changes in their clinical reasoning over the course of the clerkship year. In order to accomplish this, we decided to determine if the use of scoring rubrics could change the traditional subjective assessment to a more objective evaluation.

BASIC PROCEDURES: Two groups of students, one at the beginning of clerkship (Juniors) and one at the end of clerkship (Seniors), were chosen. Each group was given the same virtual patient case, a clinical scenario based on the Advanced Trauma Life Support (ATLS) Primary Trauma Survey, which had to be completed during their trauma rotation. The learner was required to make several key patient management choices based on their clinical reasoning, which would take them along different routes through the case. At the end of the case they had to create a summary report akin to sign-off. These summaries were graded independently by two domain "Experts" using a traditional subjective surgical approach to assessment and by two "Non-Experts" using two internally validated scoring rubrics. One rubric assessed procedural or domain knowledge (Procedural Rubric), while the other rubric highlighted semantic qualifiers (Semantic Rubric). Each of the rubrics was designed to reflect established components of clinical reasoning. Student's *t*-tests were used to compare the rubric scores for the two groups and Cohen's *d* was used to determine effect size. Kendall's τ was used to compare the difference between the two groups based on the "Expert's" subjective assessment. Inter-rater reliability (IRR) was determined using Cronbach's alpha.

MAIN FINDINGS: The Seniors did better than the Juniors with respect to "Procedural" issues but not for "Semantic" issues using the rubrics as assessed by the "Non-Experts". The average Procedural rubric score for the Senior group was 59% ± 13% while for the junior group, it was 51% ± 12% ($t_{(80)}$= 2.715; p = 0.008; Cohen's d = 1.53). The average Semantic rubric score for the Senior group was 31% ± 15% while for the Junior group, it was 28% ± 14% ($t_{(80)}$ = 1.010; p = .316, ns). There was no statistical difference in the marks given to the Senior versus Junior groups by the "Experts" (Kendall's τ = 0.182, p = 0.07).

Correspondence: Inquiries to Nancy Posel, Faculty of Medicine, McGill University, Suite 301, 4920 de Maisonneuve, Montreal, Que., Canada H3Z 1N1; e-mail: nancy.posel@mcgill.ca

[1] Royal Victoria Hospital, 1001 avenue Décarie, Room D2-7115, Montréal, QC. Canada H4A 3J1.

[2] Faculty of Education, McGill University, 3700 McTavish Street, Montréal, QC. Canada H3A 1Y2.

[3] Montreal General Hospital, 1650 Cedar Avenue, room L9-509, Montréal, QC. Canada H3G 1A4.

[4] Jewish General Hospital Pavilion A-515 3755 Côte-Sainte-Catherine Montréal, QC, H3T 1E2.

  1931-7204/$30.00 http://dx.doi.org/10.1016/j.jsurg.2017.08.024

The IRR between the "Non-Experts" using the rubrics was higher than the IRR of the "Experts" using the traditional surgical approach to assessment. The Cronbach's alpha for the Procedural and Semantic rubrics was 0.94 and 0.97, respectively, indicating very high IRR.
The correlation between the Procedural rubric scores and "Experts" assessment was approximately $r = 0.78$, and that between the Semantic rubric and the "Experts" assessment was roughly $r = 0.66$, indicating high concurrent validity for the Procedural rubric and moderately high validity for the Semantic rubric.

PRINCIPLE CONCLUSION: Clinical reasoning, as measured by some of its "procedural" features, improves over the course of the clerkship year. Rubrics can be created to objectively assess the summary statement of an online interactive trauma VP for "procedural" issues but not for "semantic" issues. Using IRR as a measure, the quality of assessment is improved using the rubrics. The "Procedural" rubric appears to measure changes in clinical reasoning over the course of 3rd-year undergraduate clinical studies. (J Surg Ed ▮:▮▮▮-▮▮▮. © 2017 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: clinical reasoning, assessment and evaluation, virtual patient, undergraduate medical education, scoring rubric

COMPETENCIES: Medical Knowledge, Practice Based Learning, Improvement

INTRODUCTION

Clinical reasoning is "a cornerstone of medical practice."[1] The development of clinical reasoning skills permits medical learners to (a) critically analyze patient situations, (b) determine the significance of symptoms, signs, laboratory results and imaging, (c) engage in the development of a hypothesis, or differential diagnosis, (d) participate in active problem solving, and (e) define associated treatment-planning that leads to problem resolution and positive patient outcomes.[2–4] Unfortunately, the direct assessment of clinical reasoning is challenging and particularly difficult during the clerkship year,[5,6] which is especially troubling as many educators anticipate the shift toward competency-based education. At our institution, particularly in the field of trauma, students report a highly variable exposure, few opportunities for independent management of patients, and even fewer opportunities for direct observation and evaluation by supervisors.

In their "Recommendations for Clinical Skills Curricula for Undergraduate Medical Education," the American Association of Medical Colleges suggested that diverse instructional methodologies, including simulated clinical learning opportunities can and should be used to support the development of clinical reasoning skills.[7] VPs are mid-level fidelity simulated online clinical learning opportunities, also described as "multimedia, screen-based interactive patient scenarios"[1] (page 1217) nested in decision-based learning applications that can be used to teach and assess students.[8,9,10,11,12]

Our institution developed a series of trauma VPs based on the ATLS principles of the American College of Surgeons. The VPs include multiple assessment features that can parallel key features examinations, including multiple choice, inquiry assessment (multiple right and wrong answers), and summary statements, simulating "sign-off" to an attending physician. This exercise requires the student to recall the key components of the case, to organize them appropriately and to prioritize treatment procedures creating a clear and contextual analysis of the critical components associated with the VP case. Critical indicators associated with clinical reasoning emphasize learners' ability to synthesize, organize, and prioritize their clinical narrative within a comprehensive, succinct, and clear construct. The summary statement, as an instructional strategy, aligns with these indicators associated with clinical reasoning.[11,13,14]

The summary statement is initiated by an instruction to create the summary as though the student was presenting the case to an attending surgeon. Underneath this instruction is a text input field where the student types their summary. A submit button signifies completion of the summary and the application emails the summary to the instructor. The application then reveals an expert's summary to which the student can compare their own summary.

The Expert Summary

"This is (Name of Student) calling. I am a 3rd-year medical student. I (we) have a young adult male in the trauma bay who as the driver was involved in a T-bone car crash with substantial intrusion where he was sitting. There were no airbags. He smells of alcohol. He was staggering around the accident site when the ambulance arrived, but subsequently became unresponsive. At 3:15, on arrival in the trauma bay, he was somnolent and in severe respiratory distress. He has been intubated. He had a hemopneumothorax and has had a left chest tube inserted. Approximately 600 cc of blood were drained. There are multiple rib fractures and subcutaneous crepitus but he is oxygenating well at this point with a Sat of 90%. He arrived with a pressure of 90 and a pulse of 130 and RR of 30 and shallow. After the chest tube insertion and after approximately 2 L of fluid his pressure is now 100 and his pulse 120. He has been cross-matched for 6 units and we will give him blood as soon as it arrives. He has a closed head injury that has progressed since the accident. His GCS at 4:15 is 5t (E1 V1 M3). Neurosurgery has been called. On exposing and logrolling the patient there is no obvious abdominal injury, however, his left leg is shortened and internally rotated suggesting a left hip fracture. We are about to start the secondary survey and assuming he continues to stabilize we will take him to CT for a total body scan. Is there anything else you would like us to do for the moment?"