COMPETENCY BASED MEDICAL EDUCATION: O-SCORE CHARACTERISTICS OF PROCEDURAL AND COGNITIVE ASSESSMENTS IN GASTROENTEROLOGY RESIDENCY TRAINING

J. Cooper[1], M. Gozdzik[1], H. Lai[2], J. Silverman[3], K.I. Kroeker[1]

1) Department of Medicine, Division of Gastroenterology, University of Alberta, Alberta.
2) 2. Department of Dentistry & Dental Hygiene, University of Alberta, AB, Canada
3) Division of Pediatric Gastroenterology and Nutrition, University of Alberta, AB, Canada

JOINT 2024 CDDW-CLM CONFERENCE
FEBRUARY 27 – MARCH 3 · 2024 · TORONTO · CANADA

Introduction

- Competency based medical education is the new standard for medical education; implemented in Canada as 'Competence by Design' (CBD).
- CBD assesses a physician trainee's ability to demonstrate competence in CanMEDS roles via entrustable professional activities (EPAs).[1-3]
- EPAs utilize the O-SCORE originally validated for surgical/procedural evaluations, but is currently coopted for both procedural and non-procedural (cognitive) EPAs in CBD.[4,5]
- Additionally, assessor expertise has been shown to have an important role in performance assessments.[6] Factors affecting evaluations include:
  - Assessor characteristics
  - Assessor perceptions of the assessment task
  - The context of the assessment
- Both the validation of O-SCORE usage for cognitive EPAs, as well as the role of assessor characteristic in EPA evaluations has yet to be investigated in the context of CBD.
- **Study objectives**:
  - Assess for differences in O-SCORE utilization between cognitive and procedural EPAs, and whether assessor characteristics are associated with trends in assessment.

Methods

- Cross-sectional study evaluating anonymized Adult Gastroenterology subspecialty EPAs completed from Jun 2019 to Jan 2023 at the University of Alberta.
- Extracted variables from each EPA:
  - EPA type: Procedural vs non-procedural (cognitive)
  - EPA score (1-5)
  - Evaluator sex
  - Clinical vs academic practice
  - Advanced training expertise (I.e. Hepatology, therapeutics, etc.)
- Local score interpretations:
  - 5: denotes competence
  - 4: neutral score - may be accepted as evidence of competence at the discretion of the local competency committee.
  - 1-3: indicates competence was not yet achieved
- T-tests and ANOVA with post hoc Games-Howell testing were performed with 95% confidence intervals (CI). A p-value of <0.05 was significant.

Table 1. Number and proportion EPAs stratified by type and competence evaluation.

	Competence Evaluation	Number	%
Total	Achieved	1852	69.7
	Neutral	613	23.0
	Not Achieved	195	7.3
Cognitive	Achieved	1041	75.2
	Neutral	282	20.3
	Not Achieved	63	5.5
Procedural	Achieved	811	63.6
	Neutral	332	26.0
	Not Achieved	132	10.3

Table 2. Number and proportion of each EPAs score utilized stratified by EPA type and percent of staff utilizing each score

	EPA Score	Number	Percent	% Staff Utilizing
Total	1	22	0.8	29
	2	17	0.7	27
	3	156	5.9	67
	4	613	23.0	96
	5	1852	69.6	98
Cognitive	1	0	0.0	0
	2	5	0.5	11
	3	58	4.2	35
	4	281	20.2	74
	5	1041	75.1	98
Procedural	1	21	1.8	30
	2	11	0.9	17
	3	96	8.1	57
	4	301	25.5	94
	5	753	63.7	94

Table 3. Number, mean score, and mean score difference of EPAs stratified by evaluator sex and EPA type.

	Sex	Number	Mean (SD)	Difference (95% CI)
Total	Male	1778	4.64 (0.67)	0.13 (0.07, 019)*
	Female	886	4.51 (0.77)	
Cognitive	Male	937	4.74 (0.52)	0.13 (0.06, 0.19)*
	Female	448	4.62 (0.63)	
Procedural	Male	841	4.53 (0.78)	0.12 (0.24, 0.22)*
	Female	438	4.41 (0.87)	

Table 4. Number, mean score, and mean score difference of EPAs stratified by evaluator academic vs clinical status and EPA type.

	Academic vs Clinical	Number	Mean (SD)	Difference (95% CI)
Total	Academic	1488	4.51 (0.75)	0.20 (0.15, 0.25)*
	Clinical	1176	4.71 (0.62)	
Cognitive	Academic	912	4.62 (0.61)	0.18 (0.11, 0.24)*
	Clinical	473	4.82 (0.44)	
Procedural	Academic	576	4.31 (0.89)	0.33 (0.24, 0.42)*
	Clinical	703	4.64 (0.71)	

Results

- 2264 EPAs were assessed including 1385 cognitive and 879 procedural EPAs.
- Absolute number of EPAs completed by evaluators ranged from 11 to 165 with a mean of 60 (standard deviation 40).
- Results of O-SCORE usage are summarized in Table 1 and 2:
  - Majority of EPAs indicate competence, with 20-25% neutral, and <10% did not achieve competence.
  - <1/3 of evaluators utilized a score of 1 or 2 across all EPAs, and zero evaluators utilized a score of 1 for cognitive EPAs.
- Most commonly evaluators utilized 3/5 options of the O-SCORE for total EPAs.
- Most commonly evaluators utilized 2/5 and 4/5 options of the O-SCORE for cognitive vs. procedural EPAs respectively.
- Results of demographic comparisons are outlined in Table 3 and 4.
  - Male evaluators submitted higher scores on average.
  - Clinical practice evaluators submitted higher scores on average.
  - Hepatologists submitted higher scores on average (data not shown).

Conclusions

- Across total, cognitive, and procedural EPAs there are **low rates in the utilization of the whole O-SCORE scale**.
- We highlight a discrepancy between procedural and cognitive EPAs regarding O-SCORE usage. Specifically, it was **more common to utilize a greater range of the O-SCORE when procedural EPA assessments were performed**.
- There are **small but significant differences in mean EPAs score awarded between different evaluator demographics**. Interpretation is limited by biases such as preferential assessor selection by physician trainees.
- This study provides the initial data to prompt further research into validation of the O-SCORE for use in cognitive EPA evaluations.

References

1. ten Cate O. Entrustability of professional activities and competency-based training. Med Educ. 2005;39(12):1176-7.
2. Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, et al. Competency-based medical education: theory to practice. Med Teach. 2010; 32(8):638-45.
3. Frank JR, Snell L, Sherbino J, editors. CanMEDS 2015 Physician Competency Framework. Ottawa: Royal College of Physicians and Surgeons of Canada; 2015.
4. Gofton WT, Dudek NL, Wood TJ, Balaa F, Hamstra SJ. The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): a tool to assess surgical competence. Acad Med. 2012;87(10):1401-7.
5. MacEwan MJ, Dudek NL, Wood TJ, Gofton WT. Continued Validation of the O-SCORE (Ottawa Surgical Competency Operating Room Evaluation): Use in the Simulated Environment. Teach Learn Med. 2016;28(1):72-9.
6. Berendonk C, Stalmeijer RE, Schuwirth LW. Expertise in performance assessment: assessors' perspectives. Adv Health Sci Educ Theory Pract. 2013;18(4):559-71.