Comparative Effectiveness of OSCEs, Virtual OSCEs, and Traditional Written Testing Methods in Assessing Medical School Students' Competencies: A Scoping Review Interim Report

Anne Holbrook[1,2,3]; Oswin Chang[2,4]; Munil Paul Lee[5]; Simran Lohit[4]; Alan Cheng[4]; Jia Xu[4]
Jason Profetto[6]; Simon Maxwell[7,8]; Mitchell Levine[1,2,9]; Dan Perri[1-3]; Anthony J Levinson[10]; Jill Rudkowski[3]; Heather McLeod[3];

[1]Division of Clinical Pharmacology and Toxicology, McMaster University, [2]Clinical Pharmacology & Toxicology Research, St Joseph's Healthcare Hamilton, [3]Department of Medicine, McMaster University, [4]Bachelor of Health Sciences (Honours) Program, McMaster University, [5]Schulich School of Medicine & Dentistry, University of Western Ontario, [6]Department of Family Medicine, McMaster University, [7]Clinical Pharmacology Unit, University of Edinburgh, UK, [8]British Pharmacological Society, London, UK, [9]Department of Health Research Methods, Evidence and Impact (HEI), McMaster University, [10]Division of e-Learning Innovation, McMaster University

Background

By the end of their medical school training, all physicians no matter which specialty they are heading for, must be competent to prescribe medications. Selecting and prescribing medications will be the most common therapeutic intervention they make in most careers. More than 700 million prescriptions are written annually in Canada for drugs from approximately 1100 therapeutic groups.(1) Medication errors are very common in Canada as in other countries. They usually go unrecognized and pose a serious patient safety hazard.(2-6) International estimates put drug-related adverse effects as the fourth leading cause of death.(7) Medical students and residents are especially vulnerable, with prescribing errors on 7-10% of their prescriptions.(3) Prescribing errors are generally the most serious type of medication error, contributing to an estimated 22,000 deaths and a cost of £1.6 billion per year in the U.K.'s NHS.(3, 8)

Safe prescribing is a complex activity requiring detailed knowledge of dozens of therapeutic drug families, accurate patient diagnosis and individual risk assessment, excellent communication and inter-professional collaboration skills, and professionalism. The specialty of Clinical Pharmacology & Toxicology (CPT), whose role includes training physicians in competent prescribing, are championing CPT eCurriculum supplements to medical education with associated formative and summative evaluation.

For several years, our group has led an initiative that a) provided leadership in CPT Competence by Design development, b) established linkages with international leaders in CPT education, and c) evaluated methods of assessing and improving the prescribing competency of Canadian medical students using cost-effective innovation.(9-13)

Multiple choice questions (MCQs), written answer questions, and standardized patients within Objective Structured Clinical Examinations (OSCEs) form the current backbone of medical and other health professional education assessment in Canada and internationally. However, OSCEs are very resource-intense, and written answer questions are used sparingly as they are time-consuming to mark. MCQs with single best answer have become the dominant testing method but have well-known limitations - mainly the cueing effect, testing the ability to recognize an answer rather than to generate an answer, and difficulty representing many important areas of medicine where no appropriate distractors can be created.(14) Recently, computer-readable very short answer question responses have been validated as a novel pragmatic innovation with excellent results in terms of representing actual processes in clinical practice, and testing actual case-based knowledge

and skills with better discrimination than single best answer questions.[14, 15] McMaster's newly acquired ExamSoft software (https://examsoft.com/) allows 'fill in the blank' field formats that can be used to design an electronic prescription interface. This development plus computer-read scoring could lead to a major improvement in the quality and efficiency of medical student assessment. Similarly, innovative approaches to the in-person standardized patient are necessary and possible, given current platforms including Zoom with its waiting room, recording, remote camera control, and breakout rooms which include pre-exam scenario, scenario with evaluators, and post-scenario feedback. The pandemic has vastly accelerated use in actual practice of eConsults and virtual visits, requiring rapid adjustment of training and evaluation to optimize communication, including history taking at distance, non-verbal communication, physical examinations, real time clinical reasoning, and providing counseling advice in a communication scenario.[16,17]

Advancements in virtual medical education, particularly the technical advancements, are immediately relevant to other types of post-graduate education and undergraduate education as well.

The protocol for this study was registered on OSF registries (https://osf.io/6u8fy/).

Objectives

As future medical experts, medical students need to demonstrate the competencies summarized in the CanMEDS framework (professional, communicator, collaborator, leader, health advocate, scholar, medical expert).[18] This scoping review aimed to understand the advantages and the shortcomings of OSCEs in assessing the CanMEDS roles, and the comparative effectiveness of virtual versus in-person OSCEs.

Research Questions

1. For which CanMEDS domains have OSCEs been shown to be a superior evaluation method compared to multiple choice or short/long answer written questions?
2. What are the advantages and disadvantages of virtual OSCEs (examinee, examiner, simulated patient all online and at distance from each other) compared to in-person OSCEs (all 3 in same room)?

Eligibility Criteria

Inclusion criteria are highlighted below with the PICOT(S) framework:

Question 1:

- Population: Health professional students
- Intervention: OSCE stations
- Comparator: traditional written testing methods (e.g., multiple-choice questions, short answer or written essay responses)
- Outcomes: Effectiveness at assessing the CanMEDS roles (Medical Expert (the integrating role), Communicator, Collaborator, Leader, Health Advocate, Scholar, Professional), as measured by the study itself
- Timeframe: Studies published between 1946 and 2021

- Study design/type: Studies must have a control/comparison group (e.g., RCTs, cohort studies, case-control studies). Included studies must report primary data and be written in English.

Question 2:

- Population: Health professional students
- Intervention: Virtual OSCE stations
- Comparator: In-person OSCE stations
- Outcomes: Effectiveness for assessing the CanMEDS roles, logistics, feasibility, resource intensity
- Timeframe: Studies published between 1946 and 2021
- Study design/type: Studies must have a control/comparison group (e.g., RCTs, cohort studies, case-control studies). Included studies must report primary data and be written in English.

Editorials, commentaries, and guidelines will be excluded. Conference abstracts will be considered since they may have relevant information or data.

Literature search

We conducted searches in MEDLINE to find research on the effectiveness of OSCEs compared to traditional testing methods in assessing the CanMEDS. Meanwhile, we searched MEDLINE and EMBASE regarding how virtual OSCEs compare to in-person ones. The PRISMA flow diagrams[19] for the two research questions can be found in Figure 1 and Figure 2, respectively. The search strategy (Appendix 1) was created with the assistance of librarians at McMaster University and St. Joseph's Healthcare Hamilton.

No date limits were put in place. Databases were searched from inception to August 2021.

Study selection

Title and abstract screening

Screening was performed in duplicate and independently on the online platform Covidence (https://www.covidence.org/). Disagreements in screening were resolved with a consensus. Studies that passed screening with the inclusion criteria underwent full-text screening in Covidence (Appendix 2).

Full-text screening

Full-text screening was conducted in the same manner as the title and abstract screening. Exclusion rationale for full-text articles was also be recorded and reported.

Data extraction

Paired and independent data extraction was completed using Microsoft Excel. A standardized form was used by the team to extract data from a small number of studies.

Information extracted included study characteristics (e.g., publication year, study design, setting, sample size), participant characteristics (e.g., health profession, level of education/training), CanMEDS domains assessed, and benefits/drawbacks of the specific form of the OSCE.

A risk of bias assessment was not necessary because this study is a scoping review.[20]

Data synthesis

Data summary and reporting was done in accordance with the PRISMA-ScR checklist to address the two research questions.[21] The analyses are descriptive and report 1) which domains OSCEs are more/less effective at assessing than traditional testing methods, and 2) the advantages and disadvantages of virtual OSCEs compared to in-person ones.

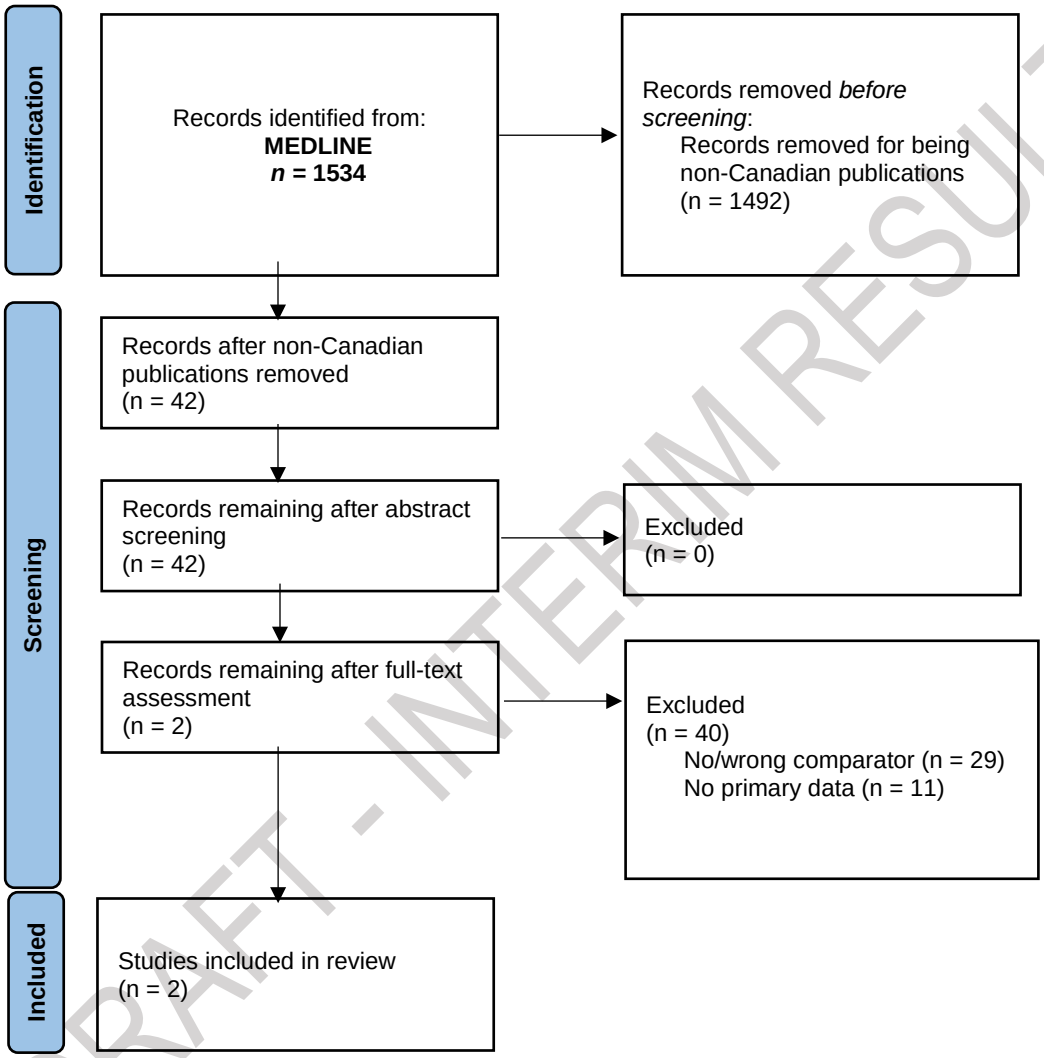

Figure 1: PRISMA flowchart for the identification and selection of studies comparing OSCEs with traditional testing methods.

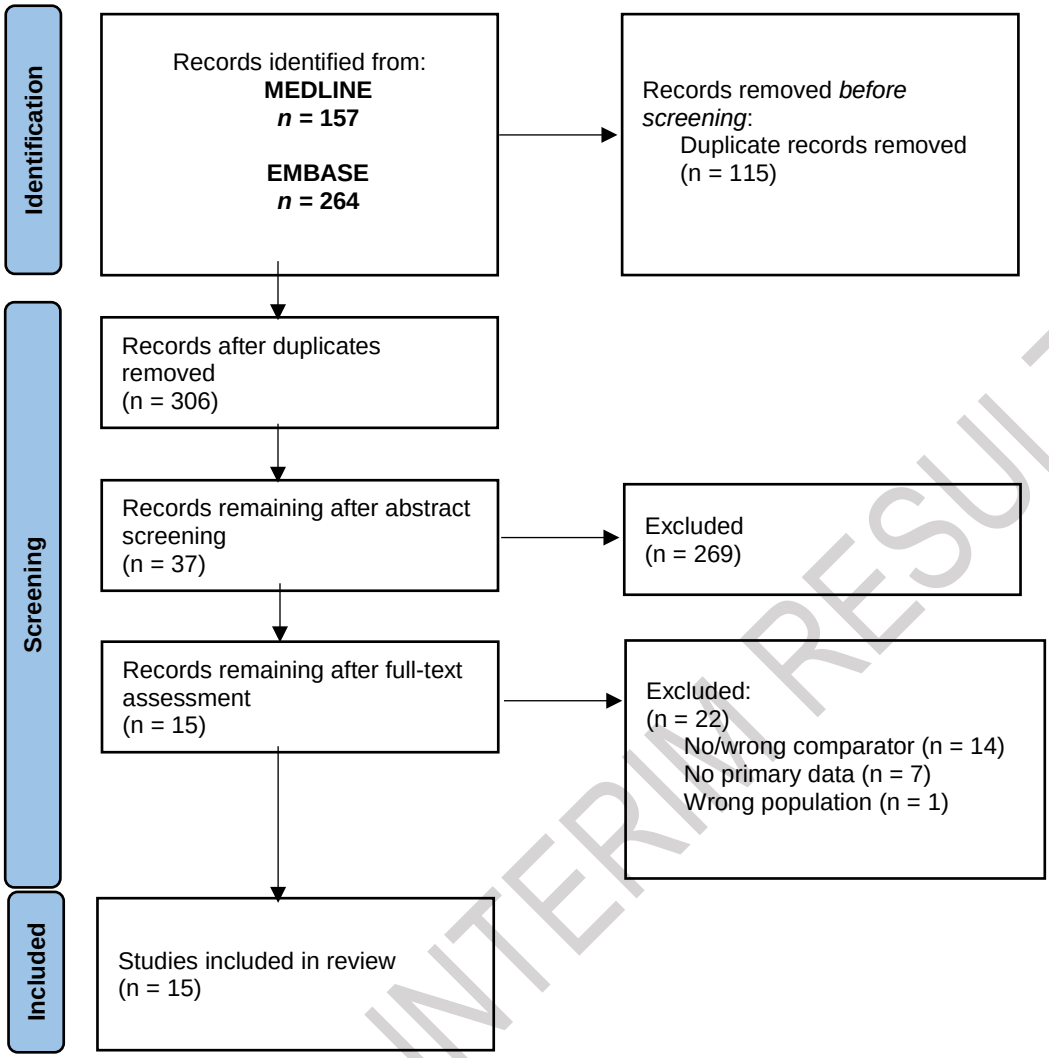

Figure 2: PRISMA flowchart for the identification and selection of studies comparing in-person and virtual OSCEs.

Results

Research Question 1

A MEDLINE search yielded 1534 citations and when restricted to Canadian publications, 42 potentially relevant articles remained. After title and abstract screening of the subset of Canadian studies, 39 of them were excluded based on the following criteria: study is not in English, study does not provide primary data, population does not include health professional students, study does not assess CanMEDS domains (or equivalent competency measures), or study does not compare OSCEs with traditional written examinations (Appendix 2). Of the 39 studies, 11 studies did not provide primary data and 29 studies did not compare OSCEs with written examinations. The remaining 2 studies then underwent full-text screening, but did not exactly meet the criteria for inclusion. These studies, however, did provide some valuable insight.

Raj et al. (2006) developed two OSCE stations for third and fifth year medical students for the assessment of core hand and knee examination skills. These stations were evaluated for content,

construct and concurrent validity, as well as for reliability. Comparisons were made with the performance of six specialists in rheumatology, as well as with the final examination scores of the students. This study was initially included in abstract screening as it was unclear which final examination was the comparator – however, this was found to also be an OSCE on full text screen. In short, while the study did demonstrate validity and reliability of its OSCE stations, there was no comparison to written examination.

Dwyer et al. (2013) developed an OSCE for orthopedic residents in years 1, 3, and 5 at the University of Toronto. The aim was to test the 6 intrinsic CanMEDS roles – i.e., all of the roles aside from Medical Expert. OSCE scores were compared to in-training evaluation reports (ITER) from the previous 12 months and an ordinal ranking made by the program directors. There was no written examination for comparison. Their study found a statistically significant difference between the scores of residents in years 1, 3, and 5, suggesting that the OSCE could differentiate between residents at different levels of training. There was good correlation between OSCE scores and program director scores, but no correlation was seen between ITER and OSCE scores.

Research Question 2

A MEDLINE and EMBASE search yielded 306 potentially relevant articles with 269 articles being removed during title and abstract screening. The remaining 37 articles underwent full text review, and 22 were excluded based on the following criteria: the study is not in English, study does not provide primary data, population does not include health professional students, study does not assess CanMEDS domains (or equivalent competency measures), logistics, feasibility, resource intensity, or other aspects, or study does not compare in-person OSCEs with virtual OSCEs (Appendix 2). From the excluded studies during full text review, 14 had no/wrong comparator, 7 had no primary data, and 1 focused on the wrong population (not health professionals).

From the 15 included studies, 10 were conducted in the United States, and 12 used a before-and-after study design. None of the papers had outcome assessors who were blinded. 11 studies focused on the field of medicine, and 12 studies examined test outcomes in students. Full characteristics of included studies can be found in Table 1.

The most assessed CanMEDS roles were medical expert and communicator, with 9 studies each. None of the studies looked at the Leader or Scholar roles. Of the 15 comparisons of overall score, 12 of them found no statistically significant difference between in-person and virtual OSCEs, while 3 found virtual OSCE scores were statistically significantly higher. When breaking down scores by CanMEDS roles, most studies also found no difference between scores. However, Communication was the most varied, where 8 comparisons found no difference, 3 found in-person scores were higher, and 4 found virtual scores were higher. For the Professional role, 3 comparisons found no score differences between in-person and virtual OSCEs, while 3 found in-person scores to be higher. Overall, it appears that virtual OSCEs and in-person OSCEs are comparable for assessing CanMEDs roles. Remaining results are summarized in Table 2.

Table 1: Study Characteristics (n=15)

Characteristics	**Value: n (%)**
Publication Year	
2021	9 (60.0)

Before 2021	6 (40.0)
Country of Origin	
USA	10 (66.7)
Other (Spain, India, Germany, Israel, Indonesia)	5 (33.3)
Study Design	
Randomized crossover trial	2 (13.3)
Prospective cohort with concurrent control	1 (6.7)
Before-and-after	12 (80.0)
Number of In-person OSCE Stations*	
<4	3 (20.0)
4-6	2 (13.3)
>6	2 (13.3)
Number of Virtual OSCE Stations*	
<4	4 (26.7)
4-6	4 (26.7)
>6	4 (26.7)
Outcome assessors were blinded to group allocation during grading	0 (0.0)
Single medical school/training program	14 (93.3)
Type of Health Profession	
Medicine	11 (73.3)
Nursing	1 (6.7)
Pharmacy	3 (20.0)
Level of Training	
Student	12 (80.0)
Resident	2 (13.3)
Practicing health professional	1 (6.7)
Virtual OSCEs have examiner, examinee, and SP[§]: Yes	8 (46.7)
Total participant size at beginning of study (in-person and virtual)*	
<50	2 (13.3)
50-99	1 (6.7)
100-149	2 (13.3)
150-200	3 (20.0)
>200	5 (33.3)

*does not add up to the total number of studies (n=15), insufficient information was provided by the included studies to make a determination

[§] Studies where the Standardized Patient was also the examiner were not counted

Table 2: Summary Results

Outcome	Value
Studies assessing each CanMEDS role: n (%)	
Medical expert	9 (60.0)
Communicator	9 (60.0)

Collaborator	2 (13.3)
Leader	0 (0.0)
Health advocate	2 (13.3)
Scholar	0 (0.0)
Professional	4 (26.7)
Unclear	3 (20.0)
Evaluation Scheme*: n (%)	
Checklist	7 (58.3)
Rubric	2 (16.7)
Other	3 (25.0)
Overall Score: n comparisons	
No statistically significant differences between virtual and in-person	12 (80.0)
In-person statistically significantly higher	0 (0.0)
Virtual statistically significantly higher	3 (20.0)
Subscale Score – Medical Expert: n comparisons	
No statistically significant differences between virtual and in-person	8 (66.7)
In-person statistically significantly higher	3 (25.0)
Virtual statistically significantly higher	2 (8.3)
Subscale Score – Communication: n comparisons	
No statistically significant differences between virtual and in-person	8 (53.3)
In-person statistically significantly higher	3 (20.0)
Virtual statistically significantly higher	4 (26.7)
Subscale Score – Collaborator: n comparisons	
No statistically significant differences between virtual and in-person	1 (100.0)
In-person statistically significantly higher	0 (0.0)
Virtual statistically significantly higher	0 (0.0)
Subscale Score – Health Advocate: n comparisons	
No statistically significant differences between virtual and in-person	2 (100.0)
In-person statistically significantly higher	0 (0.0)
Virtual statistically significantly higher	0 (0.0)
Subscale Score – Professional: n comparisons	
No statistically significant differences between virtual and in-person	3 (50.0)
In-person statistically significantly higher	3 (50.0)
Virtual statistically significantly higher	0 (0.0)

*Does not add up to the total number of studies (n=15), insufficient information was provided by the included studies to make a determination

Discussion

OSCEs are thought to be a more valid method of assessing communication skills in particular, elements of professionalism (for example, demeanor), and more complex tasks such as prescription writing and review. However, for the first research question examining comparisons between OSCEs and written exams, there were no studies found comparing OSCEs to written exams for their ability to assess for CANMEDs roles. There were several studies that investigated the validity and reliability of OSCE stations, but written examinations were not in the comparator group. Studies more often compared OSCE scores of learners at different stages of training or compared OSCEs to subjective measures of performance (such as ordinal rankings or ITERs). This finding was surprising considering the importance of OSCEs in medical education. However, since our search was restricted to Canadian studies, there is a possibility we may have missed studies outside of this scope. We will screen studies that were captured with the search strategy but were not Canadian studies to ensure a more comprehensive examination of the literature. For now, further research is still required to assess the value of OSCE examinations, and to determine whether they are more, less or equally effective as written examinations in evaluating for CANMEDs competencies.

For the second research question comparing in-person and virtual OSCEs, the overall consensus from the 15 included studies indicated that virtual OSCEs were comparable to in-person ones in its ability to assess the CanMEDS roles. This finding is reassuring considering the negative impact that the pandemic has had on in-person learning and education. Virtual OSCEs seem to be a viable alternative for in-person ones for the time being as well as for the future. However, it should be noted that are advantages and disadvantages that are unique to both formats of the examination whether it be technical difficulties for virtual OSCEs or social distancing concerns for in-person OSCEs. While we were hoping to also investigate more advantages and disadvantages like costs, logistics, or feasibility, studies either did not mention these facets or did not provide data that could be analyzed. Future explorations should consider the quality of the studies as the definition of an OSCE was not always consistent across studies along with evaluation schemes.

References

1. Canada Health Infoway. Current prescribing and dispensing landscape in Canada: CHI; 2017 [Available from: https://infocentral.infoway-inforoute.ca/en/resources/docs/med-mgmt/1778-report-2017-current-prescribing-and-dispensing-landscape-in-canada. Accessed on 23 Apr 2019].
2. Baker GR, Norton PG, Flintoft V, Blais R, Brown A, Cox J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 2004;170(11):1678-86.
3. Dornan T, Ashcroft D, Heathfield H, Lewis P, Miles J, Taylor D, et al. An in-depth investigation into causes of prescribing errors by foundation trainees in relation to their medical education: EQUIP study. London: General Medical Council. 2009:1-215.
4. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. Jama. 1998;279(15):1200-5.
5. Ryan C, Davey P, Francis J, Johnston M, Ker J, Lee AJ, et al. The Prevalence of Prescribing Errors amongst Junior Doctors in Scotland. Basic and Clinical Pharmacology and Toxicology. 2011;109:35.
6. Young H. Lack of pharmacological training causes overuse and misuse of drugs. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 2008;178(3):276.
7. Light D. New Prescription Drugs: A Major Health Risk With Few Offsetting Advantages Harvard University: Edmond J. Safra Center for Ethics; 2014 [Available from: https://ethics.harvard.edu/blog/new-prescription-drugs-major-health-risk-few-offsetting-advantages. Accessed on 23 Apr 2019].
8. Elliott RA, Camacho E, Campbell F, Jankovic D, James MMS, Kaltenthaler E, et al. Prevalence and Economic Burden of Medication Errors in The NHS in England. Rapid evidence synthesis and economic analysis of the prevalence and burden of medication error in the UK Universities of Sheffield and York: Policy Research Unit in Economic Evaluation of Health & Care Interventions (EEPRU); 2018 [Available from: http://www.eepru.org.uk/wp-content/uploads/2018/02/eepru-report-medication-error-feb-2018.pdf. Accessed on 23 Apr 2019].
9. Liu J, Wong S, Foster G, Holbrook A. Prescribing Competency of Medical Students: National Survey of Medical Education Leaders. Journal of population therapeutics and clinical pharmacology = Journal de la therapeutique des populations et de la pharamcologie clinique. 2018;25(1):e18-e24.
10. Qayyum F, Wright M, Lee M, Leung C, Sada A, Holbrook A. Medical student opinions on their training in clinical pharmacology and therapeutics. McMaster University Medical Journal. 2012;9(1):4-8.
11. Wu V, Chan O, Maxwell SR, Levine MA, Perri D, Sebalt RJ, et al. Development and Validation of the McMaster Prescribing Competency Assessment for Medical Trainees (MacPCA). Journal of population therapeutics and clinical pharmacology = Journal de la therapeutique des populations et de la pharamcologie clinique. 2015;22(2):e173-8.
12. Holbrook A, Liu JT, Rieder M, Gibson M, Levine M, Foster G, et al. Prescribing competency assessment for Canadian medical students: a pilot evaluation. Canadian medical education journal. 2019;10(1):e103-e10.

13. Zhang XY, Holbrook AM, Nguyen L, Lee J, Qahtani SA, Garcia MC, et al. Evaluation of Online Clinical Pharmacology and Toxicology Curriculum Resources for Medical Students. Manuscript under revision at British Journal of Clinical Pharmacology. 2019.
14. Sam AH, Hameed S, Harris J, Meeran K. Validity of very short answer versus single best answer questions for undergraduate assessment. BMC Med Educ. 2016;16(1):266.
15. Sam AH, Field SM, Collares CF, van der Vleuten CPM, Wass VJ, Melville C, et al. Very-short-answer questions: reliability, discrimination and acceptability. Med Educ. 2018;52(4):447-55.
16. Lucas C, Forrest G. Virtual OSCEs – Challenges and Considerations for Pharmacy Education? 2020 January 18, 2021. Available from: https://cptlpulses.com/2020/06/18/virtual-osces/.
17. Traba C, Holland B, Laboy MC, Lamba S, Chen S. A Multi-Modal Remote Clinical Skills Mini-Course Utilizing a Teaching TeleOSCE. Medical science educator. 2021:1-7.
18. CanMEDS: Better standards, better physicians, better care. Royal College of Physicians and Surgeons of Canada. July. 23, 2021. Available from: https://www.royalcollege.ca/rcsite/canmeds/canmeds-framework-e.
19. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
20. Peters MD GC, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. International journal of evidence-based healthcare. 2015;13(3):141-6.
21. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. 2018;169(7):467-73.
22. Raj N, Badcock LJ, Brown GA, Deighton CM, O'Reilly SC. Design and validation of 2 objective structured clinical examination stations to assess core undergraduate examination skills of the hand and knee. The Journal of rheumatology. 2006;34(2):421-24.
23. Dwyer T, Takahashi SG, Hynes MK, Herold J, Wasserstein D, Nousiainen M, Ferguson P, Wadey V, Murnaghan ML, Leroux T, Semple J, Hodges B, Ogilvie-Harris D. How to assess communication, professionalism, collaboration and the other intrinsic CanMEDS roles in orthopedic residents: use of an objective structured clinical examination (OSCE). Canadian journal of surgery. 2014;57(4):230-236.

Appendix 1. Search Strategies

A. Search strategy for MEDLINE

1. CanMEDS.mp.
2. Communication/
3. exp *Professional Competence/
4. Cooperative Behavior/
5. *Leadership/
6. *Patient Advocacy/
7. *Research Personnel/
8. professional*.ti,kf.
9. communicat*.ti,kf.
10. collaborat*.ti,kf.
11. leader*.ti,kf.
12. advocat*.ti,kf.
13. scholar?.ti,kf.
14. expert?.ti,kf.
15. ((clinical* OR medical* OR physician* OR professional*) ADJ2 competenc*).ti,kf.
16. or/1-15

17. (objective* ADJ3 clinical* ADJ3 exam*).ti,kf.
18. (objective structured clinical exam* OR OSCE?).mp.
19. or/17,18

20. *Educational Measurement/
21. exp *Academic Performance/
22. Test Taking Skills/
23. (test? OR testing OR exam* OR evaluat* OR assessment*).tw,kf.
24. (multiplechoice? OR multiple-choice? OR written).mp.
25. or/20-24

26. (virtual* OR online* OR on-line* OR web-base* OR digital* OR remote* OR tele*).ti,kf.
27. Education, distance/
28. or/26,27

29. Canada.cp

30. and/16,19,25,29 **(Research Question 1)**
31. and/19,28 **(Research Question 2)**

B. Search strategy for EMBASE

1. (objective* ADJ3 clinical* ADJ3 exam*).ti,kw.
2. (objective structured clinical exam* OR OSCE?).mp.
3. or/1,2

4. (virtual* OR online* OR on-line* OR web-base* OR digital* OR remote* OR tele*).ti,kw.
5. Education, distance.mp. OR educational measurement.mp.
6. or/4,5

7. and/3,6 **(Research Question 2)**

Appendix 2. Screening Exclusion Criteria

A. Research Question #1 Title/Abstract Screening Exclusion Criteria

1. Is the study population health professional students?
- Yes -> Continue
- No -> Exclude

2. Does the study supply primary data?
- Yes -> Continue
- No -> Exclude

3. Does the study compare OSCEs with traditional testing methods, as above?
- Yes -> Continue
- No -> Exclude

4. Does the study measure effectiveness for assessing one or more CanMEDS domains (Medical Expert, Communicator, Collaborator, Leader, Health Advocate, Scholar, Professional)?*
- Yes -> Continue
- No -> Exclude

**Outcomes do not have to be expressly stated as being part of the CanMEDS framework. For example, a study looking at teamwork would be classified under Collaborator.*

5. Is the study written in English/is there an English summary which provides the necessary information?
- Yes -> Continue
- No -> Exclude

B. Research Question #2 Title/Abstract Screening Exclusion Criteria

1. Is the study population health professional students?
    - Yes -> Continue
    - No -> Exclude

2. Does the study supply primary data?
    - Yes -> Continue
    - No -> Exclude

3. Does the study compare in-person OSCEs with virtual OSCEs?
    - Yes -> Include
    - No -> Exclude

4. Does the study examine the effectiveness at assessing the CanMEDS roles, logistics, feasibility, resource intensity, or other aspects between the two OSCE formats?
    - Yes -> Include
    - No -> Exclude

5. Is the study written in English/is there an English summary which provides the necessary information?
    - Yes -> Continue
    - No -> Exclude